VOL. XXII AUGUST, 1938 No. 4

THE JOURNAL OF APPLIED PSYCHOLOGY

JAMES P. PORTER, *Editor*
OHIO UNIVERSITY ATHENS, OHIO

With a Board of Cooperating Editors
and
PAUL S. ACHILLES HENRY C. LINK
Associate Editors Representing
THE PSYCHOLOGICAL CORPORATION

SOME INTERESTING CONTRIBUTIONS IN THIS ISSUE:

JAY L. OTIS. The Prediction of Success in Power Sewing Machine Operating

ROY HAMLIN AND THEODORA M. ABEL. Test Pattern of Mental Defectives Skilled in Weaving

F. L. WELLS, RUTH WILLIAMS AND PHILIP FOWLER. One Hundred Superior Men

ROSS A. MCFARLAND AND CLIFFORD P. SEITZ. A Psycho-Somatic Inventory

PUBLISHED BI-MONTHLY BY THE EDITOR
OHIO UNIVERSITY, ATHENS, OHIO

PRINTED BY THE SCIENCE PRESS PRINTING COMPANY, LANCASTER, PA.

Entered as second-class matter at the post office at Athens, Ohio, under the act of March 3, 1879. Additional entry May 7, 1935, at the post office at Lancaster, Pa. Copyright 1938 by James P. Porter.

The Journal of Applied Psychology

CO-OPERATING EDITORS

W. F. Book, Indiana University, W. V. Bingham, Personnel Research Federation, New York; M. R. Brooks, State Teachers College, San Diego Calif.; H. E. Burtt, Ohio State University; Edmund H. Gault, Northwestern University; William Healy, Baker Foundation, Juvenile Court, Boston; H. L. Hollingworth, Columbia University; J. R. W. Watkins, Atlantic University; Herbert S. Langfeld, Princeton University; James H. Leuba, Bryn Mawr College; Henry M. McComas, Johns Hopkins University; Walter D. Scott, Northwestern University; C. E. Seashore, University of Iowa; Daniel Starch, New York; Edward K. Strong, Jr., Leland Stanford Jr. University; L. M. Terman, Leland Stanford Jr. University; E. L. Thorndike, Columbia University; Joseph Jastrow, New School of Social Research, New York; G. M. Whipple, Denver, Mass.; Helen Thompson Woolley, Teachers College, Columbia University.

The scope of The Journal of Applied Psychology will include the following: (a) The application of psychology to vocational activities, such as law, art, public speaking, industrial and commercial work, and problems of business appeals. (b) Studies of individual personalities, such as types of character, special talents, genius, and individual differences, including the problems of mental diagnosis and vocational prognosis. (c) The influence of general environment conditions, such as climate, weather, humidity, temperature; also such conditions as nutrition, fatigue, etc. (d) The psychology of everyday activities, such as reading, writing, speaking, singing, playing games or playing musical instruments, sports, etc.

Communications, manuscripts, reviews, summaries, notices, books and offprints for review should be addressed to Dr. James P. Porter, Ohio University, Athens, Ohio. Contributors are requested to follow the conventional method of citation employed by the Journal; in every case the year, volume and pages should be specified.

The Journal of Applied Psychology is published bi-monthly. The subscription price is $6.00 a year. Single numbers $1.25 to $3.00, depending upon the number of pages. Remittances and business communications should be addressed to the Editor, James P. Porter, Ohio University, Athens, Ohio.

Vol. XXII AUGUST, 1938 No. 4

THE JOURNAL OF APPLIED PSYCHOLOGY

Edited by
JAMES P. PORTER, Ohio University
Athens, Ohio

TABLE OF CONTENTS

ROSS A. MCFARLAND AND CLIFFORD P. SEITZ. A Psycho-Somatic Inventory 327

ROSS STAGNER. Marital Similarity in Socio-Economic Attitudes 340

SAMUEL P. HAYES, JR. A Note on Personality and Family Position 347

JAY L. OTIS. The Prediction of Success in Power Sewing Machine Operating 350

F. L. WELLS, RUTH WILLIAMS AND PHILIP FOWLER. One Hundred Superior Men 367

ROY HAMLIN AND THEODORA M. ABEL. Test Pattern of Mental Defectives Skilled in Weaving 385

GEORGE K. BENNETT. A Simplified Scoring Method for the Bernreuter Personality Inventory 390

EMMETT L. SCHOTT. Superior Intelligence in Patients with Fröhlich's Syndrome 395

MAPHEUS SMITH. An Experiment to Test the Reliability of Estimates of Use of Time 400

RALPH R. BROWN AND VICTOR H. VOGEL. Psycho-physiological Reactions Following Painful Stimuli under Hypnotic Analgesia Contrasted with Gas Anaesthesia and Novocain Block 408

DONALD G. PATERSON AND MILES A. TINKER. The Part-Whole Proportion Illusion 421

JOSEPH S. KOPAS. The Point-Tally: A Modified Method of Scoring the Strong's Vocational Interest Test 426

Brief Reports of Mental Testing:

HERMAN G. CANADY. Sex Differences in Intelligence Among Negro College Freshmen 437

News and Notes 440

Book Reviews 445

New Books and Pamphlets 453

A PSYCHO-SOMATIC INVENTORY[1]

ROSS A. McFARLAND AND CLIFFORD P. SEITZ[2]

INTRODUCTION

THE purpose of this study was to develop a relatively short questionnaire which might prove to be of value in distinguishing psychoneurotic from normal individuals. The investigation was planned so as to proceed empirically, by obtaining the most frequent complaints of patients from their clinical histories and interviews rather than logically, that is, in listing items based upon current concepts of the psychoneuroses. These items were then roughly grouped as psychological and as physiological so as to analyze the relationships existing between those two inseparable variables.

In the contemporary literature of the biological sciences it is becoming increasingly clear that the lines which frequently have been drawn between "psychic and somatic," "mental and morphological," and "functional and organic" are quite illusory. The investigations initiated by Pavlov (12) and Cannon (2) have demonstrated how "mental" and emotional reactions are intimately associated with extensive organic correlates throughout the body. It is also clearly recognized that an alteration of the organic constants in the internal environment, such as in the oxygen saturation of the arterial blood (1, 10) or body temperature (6), either in clinical disorders or under controlled laboratory experiments, will bring about

[1] Published by Psychological Corporation, 522 Fifth Avenue, New York City.

[2] From the Department of Psychology, Columbia University. The authors wish to express their indebtedness to the Director of the Neurological Institute, New York, for permission to use its case material. We are especially indebted to Jas. H. Huddleson, M.D., for the selection of the patients in the Vanderbilt Clinic.

changes in mental and emotional behavior. One need only consult a review of the literature such as that of Dunbar (3) or Fetterman (4) to find numerous illustrations, both clinical and experimental, of the psycho-somatic continuum. It appeared to be of interest, therefore, to group the items as indicated so as to observe the frequency with which a high or low score in one would be found in the other.

A complete review of the literature on neurotic inventories, as well as a discussion of the validity of the various methods employed in their development, may be found in the recent review by Mosier (11). The method can be traced to Woodworth's Personal Data Sheet (15), which was validated by using only those questions which seemed to differentiate soldiers in the World War known to be mentally ill from a group of normal subjects. In comparing the scores obtained from college students with those obtained from a group of psychotic and neurotic patients, Garrett and Schneck (5) found that many of Woodworth's items were of little diagnostic value. More recently, the studies of Landis *et al.* (7), (8), (9) have shown that even schizophrenic patients did not mark items which were agreed upon by psychiatrists to be symptomatic of the disorder any more frequently than did the normal subjects. These and other studies have given rise to skepticism concerning the clinical method of validation. Furthermore, it has been demonstrated that validation by other methods, such as those of internal consistency, or one test against another, has all the disadvantages inherent in the clinical method, if not additional ones (cf. Mosier (11)).

Most of the difficulty in devising an inventory which has a high degree of validity in differentiating the mentally ill from the normal is related to the unreliability of clinical diagnosis within the field of psychiatry itself, as well as to the transitory nature of different traits of each patient during various stages of his disorder. Since there appears to be no clear-cut clinical entity of the psychoneurosis, to say nothing of the various classifications within this general set of symp-

toms, such as anxiety state, hysteria or neurasthenia, it is obviously difficult to develop a test relating to a syndrome with such a non-specific etiology and clinical picture. For these reasons we prefer to consider the term "psychoneurotic" as descriptive of certain kinds of behavior deviating from the so-called normal in terms of frequency of complaints of alterations in behavior rather than as a clinical entity. The use of the term "psychoneurotic" also implies that the symptoms are determined by "mental" and emotional factors more than by organic ones. However, tubercular or cardiac patients might make equally low scores on the psycho-somatic inventory as the neurotics due to altered organic functions.

Procedure

The development of this inventory differs from most others in that the "normality" or "abnormality" of each item was established empirically by a statistical analysis of the responses of two groups, one known to be abnormal (psychoneurotics from an outpatient clinic) and the other supposedly normal (college students).

The two lists of physiological and psychological items consisting of 54 questions in each group were obtained from the case histories of patients visiting the clinic who had been diagnosed as psychoneurotic. No case histories were analyzed which had any known organic illness. These lists were then compared with a number of standardized neurotic inventories. Several items were added which had been found by other investigators to be highly reliable or diagnostic of psychoneurotic traits. In order to avoid a possible concentration of responses in what the subjects might consider the middle or average response, four categories were used instead of the usual three; that is, *often, at times, seldom,* and *never.*

The questions were submitted to 100 male and 150 female neurotics and 550 male and 166 female normal subjects. The psychoneurotics were outpatients from the Psychiatric and Neurological Institutes of the Columbia Medical Center. The

normal subjects were, for the most part, college students at Columbia and Duke Universities. The modal age of the latter group was 24; of the psychoneurotic group, 28 (male) and 30 (female). The range for the patients was considerably greater than in the case of the control group.

The education and the economic level of the two groups were quite different; that is, a college population compared with outpatients from a free clinic. However, more than 50% of the neurotics included persons with high school and college experience. The test was not given to patients who had no education and were unable to read and interpret the questions. The groups were similar in that most of the subjects were selected from the population of a large city.

The patients took the test while they were waiting in the clinic for a consultation with a physician. In many cases they answered the questions on the inventory just after taking the Schneider Index of neurocirculatory efficiency.[3] They took the matter seriously, since most of them felt the need of medical advice and many of the questions appeared to be related to their illnesses. Some of the items were answered hesitantly, but without assistance, except in so far as to explain the meaning of certain terminology. On the average, the patients took from two to three times as long to fill out the inventory as did the normal subjects. The test was given to the control group in classes of general and abnormal psychology. They responded promptly and directly, manifesting considerable interest in the test.

Before the statistical treatment of the data was begun, the inventory of each patient was checked with the case history to observe whether the diagnosis had been changed during the course of the consultations at the clinic. At least one-third of the inventories were eliminated because the illness turned

[3] This study has been reported separately under the title "Neurocirculatory Reactions in the Psychoneuroses Studied by the Schneider Method." Ross A. McFarland, and James H. Huddleson. American Jr. Psychiatry, Vol. 93, November, 1936.

out to be one of the major psychoses or because of some organic disorder. The main classifications into which both the male and female psychoneurotic fell were: (1) anxiety state; (2) anxiety hysteria; (3) neurasthenia; (4) conversion hysteria; (5) psychasthenia; and (6) mixed psychoneuroses. The number of patients in these groups was not sufficient to treat each syndrome separately so as to determine a possible grouping of the essential characteristics of each group.

Treatment of Data. The responses of the various groups were tallied by questions and expressed in percentage units so as to avoid the difficulty of unequal groups. Each list of questions (physiological and psychological) was handled separately. Those questions were then eliminated in which the ratio between psychoneurotics and normal response was less than 2:1 on all four categories of response; that is, *often, at times, seldom,* and *never.* In the final analysis, six items were excluded in this way. Eight others were eliminated because the experimenter found them to be ambiguous or frequently misinterpreted.

TABLE 1

Sample Questions from the Inventory in Order of Reliability

PHYSIOLOGICAL QUESTIONS	PSYCHOLOGICAL QUESTIONS
Feel nervously broken down	Get irritated or upset easily
Feel well and happy	Feel just miserable
Excited or nervous (inward tension)	Things go wrong for you by no fault of your own
Feel faint or do faint	Feelings easily hurt
Spells of dizziness	Get discouraged easily
Constipation	Queer feeling as if you were not your old self
Physically depressed or miserable	Feel grouchy
Fidgety and restless	Feel afraid in many situations
Fatigued or exhausted	Worry over trifles
Difficulty in breathing	Get so discouraged cannot work properly

The weights for final scoring were based upon the data of 50 male psychoneurotics and 230 normal male subjects. The

former were only those definitely known to be neurotic. This diagnosis was based upon at least six or more examinations, consultations, and interviews with neurologists and psychiatrists at the clinic.

The percentage scores for the psychoneurotic group, arrived at as described above, were subtracted algebraically from the percentage scores for the normal group, a minus difference indicating a larger psychoneurotic response, a positive difference a greater normal response. All the found differences were then divided by the largest difference score; that is, the largest score was equal to 1.0 and the others as per cent of 1.0.

In order to determine the validity of the procedure thus far, a sample of normal and neurotic inventories, both male and female, chosen at random were scored and the significance of the differences determined. Since the female responses were not included in obtaining the weights, it was possible to compare the validity of the method not only in terms of the group upon which the weights were determined but on entirely new groups known to be neurotic (same criteria as above) and normal. The differences were found statistically significant ($D/\sigma D$ greater than 3) indicating the extent to which the test might discriminate between two such groups, as well as the adequacy of the technique of analysis (cf. Table 2).

The crude weights were then multiplied by a constant (10) and rounded off to the nearest whole number. To justify this procedure the male groups of the above samples were rescored and the reliability of the difference between the crude and final weights determined. The differences were found to be negligible (cf. Table 2).

All of the inventories, including those on which the weights were determined as well as those collected after the analysis was begun, were scored and the percentiles obtained. The percentiles are, therefore, based upon the combined data of 966 inventories.

In this study an attempt has been made to follow a middle course in the thoroughness with which the data has been

TABLE 2*

Means, Standard Deviations, D/σD for a Sample of Normal and Neurotic Male and Female Inventories Scored with Crude and Final Weights

	NUMBER OF CASES	PHYSIOLOGICAL		PSYCHOLOGICAL	
		Crude Weights	Final Weights	Crude Weights	Final Weights
Normal					
Male	41				
Mean		14.69	146.1	14.14	143.6
S.D.		2.9	35.5	4.5	34.6
Female	41				
Mean		11.42		9.32	
S.D.		4.8		5.8	
Neurotic					
Male	41				
Mean		−2.60	−27.80	−2.55	−29.09
S.D.		7.8	77.5	9.8	98.6
Female	41				
Mean		−3.57		−1.89	
S.D.		7.3		8.5	
D/σD between Normals and Neurotics					
Male		13.29		9.94	
Female		10.98		6.98	

* The scores in this table are different from those presented in Table 3 due to the fact that these represent the responses before the elimination of ambiguous questions and those which did not differentiate the neurotic from the normals.

analyzed. In many studies the statistical analysis has been extremely simple, such as arbitrarily weighting the responses 1, 2, 3, 4, or, at the other extreme, a very complicated procedure such as factor analysis. If a psychoneurotic inventory is to be of any significance, it deserves a more thorough study than the first of the methods mentioned above. On the other hand, there are so many obvious objections to the methods necessarily involved in the use of questionnaires that their reliability cannot be greatly improved by extensive statistical

analyses implied by factor analysis. Strong (13) has found, for example, that very complex procedures in personality tests of this sort give little additional information to the more simplified ones.

Results

The mean, range, sigma, and number of subjects in each group are shown in Table 3. The normal males made, on the average, higher scores than did the females. The sex difference in the experimental group is less pronounced, but in the opposite direction, that is, the female subjects tended to make more normal scores than the males. The males in both the control and experimental groups showed more variability than did the females; however, the difference is not statistically reliable, the D/σD being less than 2.

TABLE 3

Means and Variabilities for the Different Groups

	NUMBER OF CASES	RANGE	MEAN	SIGMA
Normal Males	550			
Physiological		−115 to +190	110.71	47.81
Psychological		−132 to +226	105.49	64.31
Total score		−171 to +387	216.20	105.00
Normal Females	166			
Physiological		− 51 to +192	103.02	49.12
Psychological		−130 to +223	80.08	62.87
Total score		−176 to +374	183.10	100.40
Psychoneurotic Males	100			
Physiological		−170 to +120	−32.86	68.91
Psychological		−180 to +163	−34.97	80.31
Total score		−340 to +258	−67.83	138.00
Psychoneurotic Females	150			
Physiological		−154 to +148	−27.33	66.96
Psychological		−194 to +182	−19.48	69.56
Total score		−302 to +327	−46.81	118.70

The results of the correlations between the physiological and psychological sections of the inventory, as shown below, indicate that there is a fair degree of probability that the person who scores high or low in one section of the test responds in a somewhat similar way to the other. The amount of overlapping between these two parts of the inventory was greater for the males than for the females.

	Correlation	
Normals		
Males	.75	.01
Females	.60	.03
Psychoneurotics		
Males	.71	.03
Females	.51	.04

Validity. The reliability of the observed differences between the control and experimental groups, both for the males and for the females, are shown in Table 4. The differences in the mean scores for the two groups are completely reliable statistically. The variability of the neurotics is also significantly different from the normals, although this is not so marked (cf. Table 4).

TABLE 4

Reliability of Differences between Normals and Psychoneurotics in terms of $D/\sigma D$

	PHYSIOLOGICAL	PSYCHOLOGICAL	TOTAL
Males			
$D/\sigma D$ means	19.97	16.70	19.60
$D/\sigma D$ sigmas	4.15	2.67	3.22
Females			
$D/\sigma D$ means	19.54	13.28	18.49
$D/\sigma D$ sigmas	3.78	1.26	2.08

The per cent of the male and female patients reaching or exceeding the median of the normals for their group and

similarly for the normals which reach or exceed the psychoneurotics is shown in Table 5. The amount of overlap between the two groups (both male and female) was not large (cf. Table 5).

TABLE 5

Overlapping of the Groups

	PHYSIOLOGICAL	PSYCHOLOGICAL	TOTAL
Per cent of psychoneurotics reaching or exceeding the median of the normals			
Males	0	3	0
Females	4	4	2
Per cent of normals below the median of psychoneurotics			
Males	0	2	1
Females	0	6	1
Per cent of females reaching or exceeding median of males			
Normals	46	30	36
Neurotics	55	61	56

Reliability. The split half reliability obtained on 100 normal males and stepped up by the Spearman-Brown Prophecy Formula was found to be:

.86 ..... physiological questions
.80 ..... psychological questions
.87 ..... total score

These results are somewhat similar to the reliabilities obtained in other tests of this kind; for example, the Thurstone Personality Schedule (14). The fairly low reliability may be attributed, for the most part, to the heterogeneity of the questions and the consequent differences in weights. As has been pointed out (11) permutation of the items among themselves would make possible any reliability coefficient between zero and unity.

The reliability of the inventory was also determined by the

test-retest method. The conditions of retesting were unfavorable, since the period of time during which the data was collected could not be controlled. The group was also small (52 male normals) and the interval between testing was not the same for all subjects. It was impossible to give the test a second time to the patients. The correlations were as follows:

.73 ..... physiological questions
.75 ..... psychological questions
.75 ..... total score

Interpretation of Norms

In considering any individual score a final decision regarding the classification as "psychoneurotic" largely depends on the clinical picture. It is evident that a decision as to what constitutes a neurotic in comparison to a normal response is somewhat arbitrary and largely a matter of degree. In order that a tentative measure or scale be made available, the following classification in terms of score based upon the findings of this investigation is suggested below.

	Normal	*Neurotic*
Physiological Scores		
Males	above 85	below 10
Females	above 75	below 10
Psychological Scores		
Males	above 65	below 10
Females	above 45	below 25
Total Scores		
Males	above 175	below 30
Females	above 135	below 25

Summary and Conclusions

A series of complaints, roughly classified as physiological and psychological, from the case histories and interviews of neurotic patients, has been arranged into a psycho-somatic inventory. In the final arrangement of the items only those

questions were included in which the ratio between psychoneurotics and normals was greater than two to one. The test was standardized on a group of 250 patients diagnosed as psychoneurotic after a number of consultations with psychiatrists and neurologists in the outpatient department of the Vanderbilt Clinic. Only the inventories of patients who were free from known organic disease and psychotic manifestations were used. The control group was made up of 716 college students. The validity and reliability of the inventory in differentiating the patients from the normal subjects were determined by the usual statistical procedures. A tentative classification was established on the basis of scores for this particular group of patients and control subjects.

The following conclusions may be stated tentatively concerning the findings of this investigation:

1. There was a significant difference between the psychoneurotics and the control group in answering the questions of this inventory which proved to be highly reliable statistically. Only 2 to 4 per cent of the psychoneurotics reached or exceeded the median of the normal individuals.

2. The items on the physiological and psychological parts of the questionnaire measure the same factors but to a slight extent as indicated by a 25–50 per cent overlap.

3. The reliability of the inventory as determined by the split half and test-retest method was .80–.87 (uncorrected).

4. The questionnaire method, either in differentiating psychoneurotic patients from normal subjects or in detecting subjects within the so-called normal group who have a large number of psychoneurotic traits, is of limited value because of the lack of a clear-cut clinical description of what constitutes a psychoneurotic and also because of the subjective nature of the test. An inventory such as the one developed in this study, which has been based upon empirical procedures with patients rather than logical ones, tends to differentiate patients diagnosed by the clinicians as psychoneurotic from the supposedly normal subjects. In a certain percentage of

the cases, depending upon the degree of cooperation of the subjects, such an inventory will detect the subject in the normal group who might profit by psychiatric guidance and treatment.

References

1. Babcroft, J. The Brain and Its Environment. Yale University Press, New Haven, 1938.
2. Cannon, W. B. Stresses and strains of homeostasis. *Am. Jr. Med. Sci.*, 189, 1, 1935.
3. Dunbar, H. F. Emotions and Bodily Changes. A Survey of Literature on Psychosomatic Interrelationships, 1910–1933. Columbia University Press, New York, 1935.
4. Fetterman, J. L. A correlation of psychic and somatic disorders. *J. A. M. A.*, 106, 26, 1936.
5. Garrett, H. E., and Schneck, M. R. A study of the discriminative value of the Woodworth Personal Data Sheet. *Jr. Gen. Psychol.*, 1, 459, 1928.
6. Hoagland, H. Pacemakers in Relation to Aspects of Behavior. Macmillan, New York, 1935.
7. Landis, C. The validity of certain questions which purport to measure neurotic tendencies. *Jr. Applied Psych.*, 18, 343, 1934.
8. Landis, C., Zubin, J., and Katz, S. E. Empirical evaluation of three personality adjustment inventories. *Jr. Educ. Psychol.*, May, 1935.
9. Page, J., Landis, C., and Katz, S. E. Schizophrenic traits in the functional psychoses and in normal individuals. *Am. Jr. Psychiat.*, 13, 1213, 1934.
10. McFarland, R. A. The psychological effects of oxygen deprivation (anoxemia) on human behavior. *Arch. of Psychol.* No. 145, 1932.
11. Mosier, C. I. On the validity of neurotic questionnaires. *Jr. Social Psych.*, 1, 3, 1938.
12. Pavlov, I. P. Cf. Frovlov, Y. P., Pavlov and His School. Macmillan, New York, 1938.
13. Strong, E. K. Procedure for scoring an interest test. *Psychol. Clinic*, 19, 63, 1930.
14. Thurstone, L. L., and Thurstone, T. G. A neurotic inventory. *Jr. Soc. Psychol.*, 1, 18, 1930.
15. Woodworth, R. S. Published in Hollingworth, H. L. The Personality Inventory. Psychology of Functional Neuroses. Appleton, New York, 1920.

MARITAL SIMILARITY IN SOCIO-ECONOMIC ATTITUDES

ROSS STAGNER
University of Akron

THE problem of resemblances between various members of the same family with respect to any significant psychological characteristic has long been considered important. Because of the greater availability of children, however, most of the data now recorded seem to deal with comparisons at early age levels. It seems much more important to study the degree of similarity existing between adults, or between parents and children. The present study seeks to present precise quantitative determinations of the amount of similarity between husband and wife with respect to reaction to stimulus items of a socio-economic nature.

PROCEDURE

Two groups are reported in this experiment. Group I was composed of fathers and mothers of freshmen at the University of Akron during the year 1935–36. Freshmen were given a set of forms on which they reacted to various socio-economic stimuli. They were, at the same time, given two extra forms on which it was asked that they obtain the reactions of their parents. No names were asked, but papers were matched by numbers on the test forms. The first 100 complete sets returned form the basis for the data reported for group I. Group II, studied in the same way, was composed of parents of children in one of the Akron high schools during 1936–37. Fifty couples were chosen at random from this group for the data herein reported.

The stimuli used in these test blanks were a list of 40 stereotyped terms of religious, political or economic significance,

such as "communist," "American Liberty League," etc.; ten opinions of a conservative nature, on which approval or disapproval could be registered on a five-point scale; and a set of 38 paired comparisons, nouns and adjectives of a rather random sort. The reactions of the subjects were: crossing out disliked stereotypes; approving or disapproving opinions; and indicating preference in the paired comparisons. Only the data from the first two are of interest in this connection.

The conditions under which the data were collected were not ideal. We have no way of knowing how much collaboration took place in the marking of the papers. Students were asked to tell their parents not to consult each other, but of course we do not know if this happened. We may point out, however, that a normal distribution of disagreements was found, which suggests that no systematic consultation occurred.

The treatment of the data consisted for the most part of computing tetrachoric correlations showing the precise degree of agreement between husband and wife on each specific reaction. A few other measures are reported as incidental attempts to show the meaning of the tetrachoric coefficients.

Results

The results for the forty socio-economic stereotypes are summarized in Table I. The range of coefficients is from .10 to .93 for Group I and from .08 to .95 for Group II. The median is .56 for I and .58 for II. In general, then, there is a substantial order of agreement between mates on such highly specific responses as crossing out disliked stereotypes.

We next attempted to see whether certain terms would be agreed upon in each group, by correlating the magnitudes of the coefficients for groups I and II. The order of magnitude of the coefficients for the two groups is correlated only to the extent of .30. It does not appear, therefore, that we can predict with much accuracy from group to group on which terms high agreement will be shown.

TABLE I
Agreement on Crossing Out Stereotypes

	I	II		I	II
Spiritualist	.45	.50	Compulsory military training	.46	.67
Republican	.74	.87	Birth control	.61	.80
Cigarettes	.58	.45	American Liberty League	.51	.65
Communist	.55	.69	Ku Klux Klan	.64	.40
Methodist	.93	.82	Sterilization	.61	.57
Soviet	.61	.58	Minimum wage law	.47	.52
Censorship	.32	.35	Socialist	.50	.53
Wall Street	.61	.70	Holy Communion	.93	.60
Resurrection	.31	.64	Nazi	.67	.95
Labor Union	.88	.08	Big Interests	.52	.74
Big Navy	.70	.95	Unitarian	.67	.65
Capitalist	.60	.56	Supreme Court	.51	.28
Pacifism	.48	.37	I. W. W.	.80	.54
Evolution	.66	.58	Democrat	.60	.69
Child Labor	.10	.28	Strike	.56	.42
Buddhism	.80	.65	Fascist	.58	.59
Radical	.52	.59	States' Rights	.53	.45
Divorce	.26	.54	Townsend Plan	.33	.57
Trusts	.44	.58	Profit system	.48	.50
Revolution	.26	.62	Revival meeting	.65	.70
			Median	.56	.58

Taking the ten highest coefficients in each column, we find only four agreeing: Republican, Methodist, Big Navy and Nazi. There is nothing about these terms that suggests why they would be sources of high agreement. Taking the ten lowest coefficients, we find only three coinciding: censorship, child labor and spiritualist. We might say that these are distinguished by generally lower emotional intensity, although this is not clearly indicated by the number of crossouts. "Big Navy," for example, is crossed out by less than 25% of the group, while "child labor" is marked out by over 80%.

There is a slight tendency for terms of a religious nature to have higher agreement than those of politico-economic significance. Taking such terms as birth control, Methodist, resurrection, etc., we find mean r's of .62 and .63 for the two

groups, while for the other group we find a mean of .55 in each group. Because of the large error of tetrachoric correlations, this difference may not be significant.

It may be suggested that an important determiner of the size of the correlation is the factor of sex difference. Sex difference in number crossing out each term was determined and correlated with the magnitude of the tetrachoric r for each stereotype. The rank-difference coefficient was only .20, and incidentally was in the direction of higher marital agreement for those words on which larger sex differences were observed. Thus we cannot say that the characteristic sex differences in reaction to particular terms caused the divergence from perfect agreement.

One further fact may be noted about these data. Some of the highest correlations are on words which designate organized groups: Republican, Communist, Methodist, Nazi, etc. To a certain extent, then, the uniformity of marking may be

TABLE II

Agreement on Endorsing Opinions

	I	II
The U. S. should have the largest military and naval air fleets in the world	.66	.40
The U. S. should stop immigration to give American workers more jobs	.60	.67
Recovery has been delayed by the large number of strikes	.70	.62
The U. S. should force the European countries to pay off their war debts	.62	.48
The people who complain most about the depression wouldn't take a job if you gave it to them	.63	.51
Unemployment insurance would saddle us with a nation of idlers	.38	.82
If we buy European-made goods, we make the depression in this country last longer	.43	.31
The vast majority of those in the lower economic classes are there because they are stupid or shiftless, or both	.50	.70
If we have unemployment, we should deport the excess workers in the U. S. to their native countries	.54	.62
Labor unions are justifiable only if they refrain from the use of strikes	.69	.79
Median	.61	.62

due to uniformity of group affiliation for husbands and wives.

In Table II are reported the correlations for the ten opinions used in the study. The range is somewhat shorter than for stereotypes, from .38 to .70 (group I) and from .31 to .82 (group II). Medians are .61 and .62. Incidentally it is interesting to note that the ten opinions yield a single conservatism score which correlates .60 for the husbands and wives of group I. This suggests that the total test correlation is merely a summation of the correlations of the individual items.

Again we find no tendency for items highly agreed upon in one group to be highly agreed on in another. Indeed, the rank-difference correlation for these ten opinions is −.29. Thus there is nothing about a specific opinion which explains its high agreement value.

Discussion

In a recent study (1) Newcomb and Svehla report correlations for husbands and wives on three Thurstone attitude scales, church, war and communism. These are respectively .76, .43 and .58. Our data include few stimuli which would be logically related to the scale on "attitude toward war," but it is interesting to note that in both our groups, religious stereotypes showed higher agreement than politico-economic terms, thus checking with their finding on the relative magnitude of agreement in these fields.

Newcomb and Svehla emphasize the conclusion that agreement between two members of a family is largely a function of the institutional influences which they bring to bear upon each other. It is possible that when we deal with aggregate data, such as are represented by scores on opinion scales, this influence will be shown and others will be masked, cancelled out by the random order of their operation.

It is our feeling, however, that agreement between husbands and wives is a much more personal matter than the above

conclusion would suggest. Why, for example, do we find in group I that "divorce" yields a correlation of .26 while "birth control" and "revival meeting," presumably similar in institutional background, give correlations of .61 and .65? Why (especially when we consider that groups I and II showed many psychological similarities) do correlations jump from .08 to .88 (labor union), from .31 to .64 (resurrection) and from .26 to .62 (revolution)? And, in spite of these marked divergences in reaction to specific words, why do we find such amazing agreement on mean correlation for the two groups?

The hypothesis which offers itself is that assortative mating and reciprocal influence furnish a degree of agreement between married couples which is more predictable as to *degree* than as to *kind* of agreement. Briefly, this would mean that for any given couple we would expect agreement on a certain number of topics, but not on any particular topic. The determination of this similarity in reaction would probably be found in the selection of prospective marital partners during courtship, and in mutual suggestion after marriage. That selection on the basis of attitudes does occur is shown by the data of Kelly (2). This basic personal agreement would furnish a base-line above which common institutional factors might increase agreement with regard to specific issues. However, as these data as well as several breakdowns of the Newcomb-Svehla data show, it is only when rather large groups or wide institutional differences are considered that differences in the size of correlation coefficients become marked. We therefore feel that the institutional factor is relatively the less important component of the total determinant.

Summary

Tetrachoric correlation coefficients are used to study the agreement of married couples on socio-economic issues. Since single reactions are the basis of computation, spurious effects due to aggregate scores are excluded. Under these conditions,

wide variation in agreement on specific issues is found, but remarkable consistency in average agreement is noted.

It is argued that these data show the greater importance of person-to-person influences in determining marital agreement, and make less significant the effect of common institutional influences. The latter are more readily noted when averaged or grouped data are employed.

REFERENCES

1. Newcomb, Theodore, and Svehla, George. Intra-family relationships in attitude. *Sociometry*, 1937, *1*, 180–205.
2. Kelly, E. L. A preliminary report on psychological factors in assortative mating. *Amer. Psychol. Assn.*, Minneapolis, 1937.

A NOTE ON PERSONALITY AND FAMILY POSITION

SAMUEL P. HAYES, JR.
Mt. Holyoke College

AN investigation concerning the relation of personality to family position has recently been reported by Stagner and Katzoff in which the Bernreuter Personality Inventory was administered to 430 men at the University of Wisconsin (1). The investigators found several personality differences related more or less reliably to family position, but the differences were not considered large enough to be significant.

The present report covers a similar investigation of 76 women students at Mt. Holyoke College during the first part of 1937. None of this group were freshmen. The results were so interesting that it was thought worthwhile to report them in spite of their statistical unreliability.

The subjects took the Bernreuter Personality Inventory twice, the second time being after an interval of two weeks. The mean of the two scores was changed into a percentile score for each of the six scales. The subjects were then classified into groups in terms of their family positions, and each group's mean percentile score for each scale was computed.

In order to make possible a comparison of the present results with those given by Stagner and Katzoff, the mean scores they reported were transmuted into percentile scores.[1]

Table I presents the mean percentile scores for Scales *N*, *S* (B–2), and *D* and compares them with the percentile equivalents of the mean scores reported by Stagner and Katzoff. As the groups were obviously too small to be reliable, the standard errors of the means were not calculated.

[1] Note that the author's results are given as means of percentiles, while Stagner and Katzoff's results are given here as the percentile equivalents of the mean scores given in their article. The error thus introduced is probably not great enough, however, to vitiate the tentative inferences drawn from this comparison.

TABLE I

Mean Percentile Scores on Bernreuter Personality Inventory

FAMILY POSITION	NUMBER OF CASES	SCALE N		SCALE S (B–2)		SCALE D	
		Hayes	S & K	Hayes	S & K	Hayes	S & K
All cases	76	35	53	47	53	59	49
Only children	17	37	48	50	64	58	53
Older of two	15	32	43	50	58	62	55
Younger of two	11	37	49	47	48	60	52
Oldest of three	9	25		55		68	
Middle of three	5	36		36		61	
Youngest of three	5	45		35		40	
Oldest of four or more	6	23		45		72	
Intermediate of four or more	5	28		37		61	
Youngest of four or more	3	74		35		15	
Oldest of three or more	15	24	56	51	51	70	47
Intermediate of three or more	10	32	55	37	52	61	46
Youngest of three or more	8	56	60	35	46	31	44

The internal consistency of the results is more significant than mere indications of reliability, however. Not only do they corroborate the previous results in the direction of the differences between the family positions, but the directions of the differences in all the families with two or more children were exactly alike. The more older siblings a subject had, as long as she had at least one, the more neurotic, the less self-sufficient and the less dominant she was likely to be.

The results for the only children were not as consistent with the previous study as were the others.

The men reported by Stagner and Katzoff were, on the whole, distinctly more neurotic, more self-sufficient and less dominant than the women here reported.

Table II presents the mean percentile scores for Scales *I*, *C* and *S* (F–2). Here again, the consistency of the results was striking. For those children who had at least one sibling, having older siblings meant being more introverted, less self-confident and more sociable than their older brothers and sisters.

TABLE II

Mean Percentile Scores on Bernreuter Personality Inventory

FAMILY POSITION	NUMBER OF CASES	SCALE I	SCALE C	SCALE S (F–2)
All cases	76	30	33	40
Only children	17	36	35	51
Older of two	15	26	29	38
Younger of two	11	30	33	38
Oldest of three	9	25	26	40
Middle of three	5	30	34	27
Youngest of three	5	35	50	29
Oldest of four or more	6	19	19	37
Intermediate of four or more	5	25	30	31
Youngest of four or more	3	74	74	38
Oldest of three or more	15	23	23	43
Intermediate of three or more	10	28	32	29
Youngest of three or more	8	44	59	32

No comparison of dispossessed with undispossessed siblings was made, because of the great heterogeneity within each of those categories. Although both have been dispossessed, the first of two children differs markedly from the second of three.

In summary, the evidence, although statistically unreliable, is sufficiently consistent and striking to indicate that family position may be associated with quite significant differences in personality (as measured by such a test as the Bernreuter), at least among students in a New England women's college.

REFERENCE

1. Stagner, R., and Katzoff, E. T. Personality as related to birth order and family size. *J. Applied Psychol.*, 1936, *20*, 340–346.

THE PREDICTION OF SUCCESS IN POWER SEWING MACHINE OPERATING[1]

JAY L. OTIS

University of Pennsylvania

INTRODUCTION

THE purpose of this study was to develop an adequate criterion of success for power sewing machine operators, and to investigate the validity of certain selected tests in predicting success on this job. A series of work samples closely paralleling a training course designed to train young girls for work in this trade was developed as measures of job proficiency.

Unger[2] and Treat[3, 4] worked with subnormal girls who were being trained for the garment machine trades. Grauer[5] made an intensive investigation of fifteen women electric sewing machine operators employed in a workshop for the handicapped. These studies involved the use of subjective criteria and subjects who were handicapped either mentally or physi-

[1] The writer wishes to express his gratitude to Dr. Morris S. Viteles of the University of Pennsylvania for his valuable suggestions and for the great interest shown in this study. To Miss Ruth P. Sill, Principal of the Helen Fleisher Vocational School, for her helpfulness and cooperation while collecting the data, the writer also acknowledges his indebtedness.

[2] Unger, Edna, "Vocational Training for Subnormal Girls, An Experiment in Garment Machine Operating Trades," *Journal of Personnel Research*, 5, 1926–27, pp. 243–255.

[3] Treat, Katherine, "Tests for Garment Machine Operators," *Personnel Journal*, 8, 1929–30, pp. 19–28.

[4] Treat, Katherine, "The Significance of Test Results in Predicting Efficiency in Garment Machine Operating," *Psychological Clinic*, 19, 1930, pp. 218–230.

[5] Grauer, David, "Abilities, Attitudes, and Success," *Personnel Journal*, 12, 1934, pp. 328–333.

cally. For these reasons, in planning this experiment the main emphasis was placed on the task of securing an objective measure of job success for normal subjects.

JOB DESCRIPTION

Power sewing machine operating, sometimes called garment machine operating, is a job which is found in the needlework trade. It consists in operating a sewing machine driven by electric power. The machine is so constructed that its speed of sewing varies with the amount of pressure on the foot treadle. Most work in this trade is done on a single needle machine. There are other machines, such as double needle, buttonhole, sleever, etc., but these are usually classed as special machines and the operators are called special operators. This study is concerned with all types of work on the single needle machines.

The general duties, conditions of work, and qualifications of the operator are shown in the job specification (page 352), which includes an analytic statement of ability requirements in a job psychograph, developed according to the procedure laid down by Viteles[6]: as Viteles points out, those abilities rated 4 and 5 on the job psychograph are the ones which must be primarily considered in standardizing a battery of tests. At the same time, the tests selected should measure either directly or indirectly those abilities which have been judged to be significant in the successful performance of the job. The job specification as well as the job psychograph was based upon actual job analyses made in several garment manufacturing establishments.

GROUP STUDIED

Industry provides few opportunities to test a large sample

[6] For a complete discussion of the job psychograph, see: Viteles, M. S., "Job Specifications and Diagnostic Tests of Job Competency Designed for the Auditing Division of a Street Railway Company," *Psychological Clinic*, 14, 1922; Viteles, M. S., "Job Analysis in Vocational Guidance," *Psychological Clinic*, 15, 1924, pp. 157–182; Viteles, M. S., "*Industrial Psychology*," New York, W. W. Norton Co., 1932, pp. 150–153.

of operators who work on machines of the same type and who perform the same duties. It was found that a larger sample could be obtained at the Helen Fleisher Vocational School[7] in Philadelphia than in manufacturing plants which would cooperate in an investigation of this kind.

The power operating class of the Helen Fleisher Vocational School was chosen as the experimental group. This class was composed of girls who ranged in age from 14 to 20 years, the mean age being 16. Test results were obtained for 116 girls; of that number, criterion measures were obtained for only 52 subjects. The difference in numbers was due primarily to the fact that the girls leave school to take a job in power operating in industry before completing training in the trade school.

DEVELOPMENT OF THE QUALITY CRITERION

The criterion used was based upon an analysis of the training administered to the subjects. The measures of success which were finally chosen were in the nature of work samples. The work sample requires the subject to perform all the operations of the job, or certain selected operations as they are carried through under normal working conditions. The work samples used in this study served as an attempt to sample, under standard conditions, certain elements of job performance which were important in job success.

Five work samples were used in establishing the criterion. These five represented distinct steps in the training which each

JOB SPECIFICATION—POWER SEWING MACHINE OPERATOR

NAME OF JOB: Power Sewing Machine Operator. Garment Machine Operator.

NUMBER EMPLOYED: Varies with particular plant.

Symbols: X—Essential or Generally. O—Desirable or Occasional.

DUTIES

1. Threads machine.

[7] The writer is greatly indebted to Mrs. De Long and Mrs. Klepper, instructors in power operating, for their help in the formation of the criterion.

2. Adjusts stitch regulator.
3. Fills bobbin.
4. Replaces broken needle.
5. Checks tension of thread.
6. Makes sure machine is stitching properly.
7. Oils motor.
8. Arranges supply of garment pieces to be stitched.
9. Reports faulty cutting to supervisor.
10. Sews garments according to instructions.
11. Fixes improper stitching rejected by inspector.

Operations

				Handling material	
Starting machine	X	Make various types of seams	X	Folding	X
Stopping machine	X			Creasing	X
Regulating speed	X			Tucking	X
				Shaping	X
				Stretching	X

Nature and Conditions of Work

Hours: Daily and weekly hours vary with particular plant.

Wages: Piece work, hourly rated, and weekly wage. Varies with particular plant.

Conditions				Machines Used	
Standing	—	Permanent	X	Single Needle Lock Stitch	X
Sitting	X	Temporary	—	Single Needle Chain Stitch	X
Walking	—	Routine	X		
Reaching	O	Varied	O	Special machines are usually	
Noisy	O	Special	—	operated by special operators.	
Quiet	O	Executive	—		
		Fill-in			

Time to Learn and Nature of Training

The cotton textile N. R. A. code allowed a learning period of eight weeks. The time to learn varies with the previous experience and ability of the operator; should be a good operator in six months; can become a fair operator in three months; for specific operations which are very simple a person with good ability can be trained in several days. Trade School training is valuable.

Promotion

No definite plan.

Allied Jobs

Work on special machines such as Buttonhole Machine, Sleever.

Advantages	Disadvantages
Steady work.	Requires careful application.
Clean work.	Must work at a fast rate.
Some opportunity for promotion to special machines.	Seasonal in some firms.

Remarks

The type of garment upon which the operator works determines the quality of the work; for cheap garments speed of production is essential, for better grade garments both speed and quality workmanship are demanded.

QUALIFICATIONS

General

Sex: Female preferred.
Age Limits: 18–50.
Preferred Age: 18–30.
Race and Nationality: No limitations.

Physical

Height: Not a factor.
Weight: Not a factor.
Use of both hands and feet required.

Competency

Coordination A	4
Coordination B	5
Visual discrimination	4
Space perception	4

General Intelligence

Minimum M. A. 8 yrs.

Educational

Schooling	*English*	*Mathematics*	*Special*
5th grade ___	Correspondence ___	Addition X	_____
8th grade ___	Grammar ___	Subtraction X	_____
High school ___	Spelling ___	Multiplication	_____
Technical ___	Inquiries ___	Division	_____
Business ___	Telephone ___	Percentage	_____
Correspondence ___	Speaking O	Dols. Cts.	_____
		Decimals	
		Fractions	

Temperament

Must be cooperative in working as member of group when team work is employed. Adaptability to repetitive work.

Experience

No experience is required where firms train. Trade School training, dressmaking experience is helpful. At present, most firms demand experienced operators.

JOB PSYCHOGRAPH

Job: Power Sewing Machine Operator*

	1	2	3	4	5	REMARKS
1. Energy			X			
2. Rate of Discharge			X			
3. Endurance		X				
4. Control		X				
5. Coordination A				X		
6. Coordination B					X	
7. Initiative		X				
8. Concentration			X			
9. Distribution (of attention)			X			
10. Persistence			X			
11. Alertness		X				
12. Associability		X				
13. Visual Discrimination				X		
14. Auditory Discrimination	X					
15. Tactual Discrimination	X					
16. Kinaesthetic Discrimination			X			
17. Space Perception				X		
18. Form Perception		X				
19. Accuracy			X			
20. Visual Memory		X				
21. Auditory Memory	X					
22. Kinaesthetic Memory			X			
23. Understanding		X				
24. Understanding Q	X					
25. Observation			X			
26. Planfulness		X				
27. Intelligence	X					
28. Intellect	X					
29. Judgment	X					
30. Logical Analysis	X					
31. Language Ability	X					
32. Executive	X					

* In selecting tests to predict success on this job, particular attention must be paid to four abilities which were judged to be important. These four abilities—Coordination A, Coordination B, Visual Discrimination, and Space Perception—are defined as follows:

Coordination A. This refers to the occupational requirement in the com-

girl received. A brief description of each work sample is given below.[8]

Work Sample No. 1 (Straight Stitching on Paper)

This work sample was designed to measure proficiency in straight stitching and correct spacing of stitching by sewing rows of stitches on paper, using an unthreaded machine (two trials).

Work Sample No. 2 (Straight Stitching on Material)

Two different forms of this work sample were administered: Form 1, using the selvedge edge of the material as a guide for the first row of stitching, the subject stitched lengthwise of the material; Form 2, the subject, using the same size material ($7\frac{1}{2}'' \times 24'$), stitched from selvedge to selvedge, thus having no straight edge to guide the stitching.

Work Sample No. 3 (Patches Sewed on Cambric)

This work sample involved the sewing of six cotton patches on a piece of brown cambric. Considerable skill was required to space the patches correctly without using a gauge (two trials).

Work Sample No. 4 (Bias Strips Sewed on Cambric)

Each subject was required to sew eight strips of tape, cut on the bias, to a piece of brown cambric (two trials).

Work Sample No. 5 (Dress, Size 8 Years)

The dress used in this work sample was made of cotton material, and a fairly simple pattern was followed. In addition to the dress, the

---

bination of fine muscular movements; in the establishment of finer adjustments among these movements. The importance of this ability rates high in such occupations as fine instrument making, comptometer operating, and in fine armature winding.

Coordination B. This refers to the harmonious combination of eye and muscle action demanded by the occupation. It is movement controlled by sight. It is important in such an occupation as telephone operating, in which the operator is required to get the plug into a hole of small diameter, the location of which is perceived through the eye.

Visual Discrimination. This refers to the ability required to recognize small differences in size and shape. The smallness of the difference it is necessary to recognize in a given occupation is a measure of the degree of discrimination required by the occupation.

Space Perception. This refers to the degree to which the ability to discriminate differences in distance is required by the occupation.

[8] A complete description of these work samples may be obtained from the author.

problem also included the making of a pair of bloomers. Because of the time required to complete this work sample, only one administration was given.

Work samples 1, 3, and 4 were scored for quality objectively[9] by using a ruler and a pair of dividers. The nature of the material prevented the use of an objective scoring method for work samples 2 and 5. Three raters judged the performance on work sample no. 2, and four raters were used to evaluate the quality of work sample no. 5. All of the raters were experts in judging the quality of workmanship involved on the job of power sewing machine operating.

Table 1 contains the correlations which were obtained between trials one and two of the three work samples which were scored objectively. Table 2 gives the intercorrelation coefficients of the raters' judgments of quality of workmanship on work sample no. 2. The intercorrelation coefficients of the judges' ratings of quality of sewing on the dresses are found in Table 3.

TABLE 1

Reliability Coefficients of Quality Scores for Work Samples 1, 3, and 4

NAME OF WORK SAMPLE	CORRELATION BETWEEN TRIALS 1 AND 2	R*
Straight Stitching on Paper	.67	.80
Patches Sewed on Cambric	.71	.83
Bias Strips Sewed on Cambric	.80	.89

R* The estimated reliability coefficients for both trials combined were obtained by the Spearman-Brown prophecy formula using $n = 2$.

[9] The general plan followed for each work sample was to combine the separate quality scores on each in order to get a single quality score for that work sample. The next step was to combine the quality scores on all the work samples and thus obtain the final quality criterion. By using the formula $X' = \frac{X-M}{\sigma} + 5$, where X' equals the transmuted score, X equals the raw score, M equals the mean of the distribution, σ equals the standard deviation of the distribution, and 5 equals a constant to make all transmuted scores positive in sign, the score for each item on each work sample was changed to a standard score. This procedure was followed in combining the items on each work sample, and also in combining the quality scores on the work samples.

TABLE 2
Intercorrelation Coefficients of Raters' Judgments of Quality of Work Sample No. 2 (Form 1 and Form 2)

FORM 1

Raters	1	2	3
1	——	.80	.78
2	.80	——	.80
3	.78	.80	——

FORM 2

Raters	1	2	3
1	——	.78	.78
2	.78	——	.83
3	.78	.83	——

It was impossible from the available data to arrive directly at a reliability measure for the total quality criterion. However, there seems good reason to believe that a combination of five distinct items involving a series of tasks covering a five-hour period, and each with a reliability of approximately .80, gives a total quality criterion with adequate reliability.

When the various work samples are examined, it is found that the quality measurements are predominantly objective. However, certain judgments were made and although their reliability is high, they are judgments of performance and are composed of subjective estimates. Although the criterion is not entirely objective, it is felt that the quality criterion is an

TABLE 3
Intercorrelation Coefficients of Judges' Ratings of Quality of Workmanship on Work Sample No. 5

RATERS	1	2	3	4
1	——	.78	.80	.71
2	.78	——	.79	.76
3	.80	.79	——	.64
4	.71	.76	.64	——

accurate and reliable index of the quality of workmanship of these power operators.

DEVELOPMENT OF THE SPEED CRITERION

In addition to the quality scores, the time required to complete each work sample was taken. The time scores were converted into standard scores and the standard scores were added and divided by the number of work samples to give a speed score for each individual which was used as the speed criterion.

The reliability coefficients were computed for each work sample and are found in Table 4. The time required to com-

TABLE 4

Reliability Coefficients of Work Sample Time Scores

WORK SAMPLE	r	R*
1	.61	.76
2	.66	.80
3	.75	.86
4	.79	.88

R* Estimated reliability coefficient using Spearman-Brown prophecy formula, n = 2.

plete the first two work samples was much shorter than the time required to complete work sample 3 and work sample 4. The higher reliability coefficients of the two latter work samples is probably due to the longer time required to stitch them. Here again, as in the case of the quality criterion, there seems reason to believe that the combination of items with individual reliability noted in Table 4 gives a speed criterion of sufficient reliability. This speed criterion is an objective measure of the rapidity which the subjects displayed in sewing on these work samples.

THE QUALITY-SPEED CRITERION

An attempt was made to combine these criteria, quality and speed, by adding the standard scores for each one and thus weighting them equally. However, when a careful analysis was made of these quality-speed scores, it was found that the score was not a true measure of the proficiency of the subject.

For example, one subject was the most superior worker in the group on all quality scores, but she was one of the slowest workers. When her quality score was added to her time score, the resulting measure placed her in the average group. Unfortunately, this average score was no indication of her proficiency since it hid the fact that she was excellent in respect to workmanship and it also hid her extreme slowness.

This group included individuals who were able to maintain a high degree of speed and accuracy while operating garment machines. There were others who were both slow and inaccurate, and there were many who sacrificed speed for accuracy and vice versa. The correlation between the quality criterion and the speed criterion of $-.17$ described this lack of relationship. It was felt that combining the two criterion scores tended to hide the meaning of the resulting score in that it was difficult in many cases to determine whether or not a particular score was due to proficiency in workmanship or speed.

Fortunately, the demands of this industry are such at the present time that where speed is required, less emphasis is placed on quality. It also happens that in factories where quality workmanship is demanded, there is less emphasis placed on speed. Keeping this in mind, no attempt was made to treat the quality-speed criterion.[10]

TESTS ADMINISTERED

A battery of tests used by the Vocational Adjustment Bureau[11] to predict success of subnormal girls in power sewing

[10] Experimental justification of this point of view was found in a comparison of test scores with a combined quality-speed criterion which revealed generally lower validity for a test battery with a combined quality-speed criterion than was obtained for the independent quality and speed criterion respectively. The practical implications of this finding and theoretical analysis are that a *speed battery* should be used for predicting success on those industrial jobs where speed is the essential measure of success applied by the employer, and a *quality battery* should be employed for predicting success in industrial job situations where quality of work is of prime interest to the employer.

[11] For a complete description of these tests see Treat, Katherine,

machine operating was administered to all of the subjects. The test battery which was selected by the Employment Stabilization Research Institute[12, 13] and used by Trabue as a basis for his Occupational Ability Patterns was also administered to the experimental group. To these were added the Popplereuter Tracing Test,[14] the Popplereuter Weaving Test,[15] the Minnesota Paper Form Boards A and B, and the Otis Self-Administering Test of Mental Ability, Intermediate Examination, Form B.

TEST VALIDITY

Table 5 shows the correlations which were obtained when the tests were correlated with the two criteria. Of particular interest is the correlation between the Unger-Treat battery and the Quality Criterion and the Speed Criterion. The validity of this battery is low; however, it may still be true that the Unger-Treat battery as used by them is satisfactory in distinguishing between "probable failure" and "probable success" of subnormal girls, but used with normal subjects it is not sufficiently diagnostic to indicate the degree of success which a given normal subject can attain in this occupation.

It is obvious from an inspection of Table 5 that no one of these tests can be used to predict success in this occupation. Toops' multiple-ratio correlation formula was applied to these correlation coefficients in order to select a quality test battery

---

"Tests in Garment Machine Operating," *Psychological Clinic*, 19, 1930. In a later revision, Treat found that three tests—Porteus Maze, Trimming Test, and Paper Folding—gave practically the same results as the original five tests.

[12] For a complete description of the administration of these tests see Green, Berman, Paterson, Trabue, *A Manual of Selected Occupational Tests*, Bulletins of the Employment Stabilization Research Institute, Vol. 2, No. 3, Univ. of Minn. Press, July, 1933.

[13] The Bernreuter Personality Inventory was dropped because the test was not a suitable instrument at this age and educational level.

[14] Popplereuter, W., "Die Arbeitskurve in der Eignungsprufung," *Ind. Psychot.*, 3, 1926.

[15] *Ibid.*

TABLE 5

Correlation Coefficients of Unger-Treat Tests, Minnesota Employment Stabilization Research Institute Tests, and Other Selected Tests with the Quality and Speed Criteria

TEST	QUALITY CRITERION	SPEED CRITERION
Paper Folding	.30	−.10
Porteus Maze	.16	.09
Trimming	.29	.20
Card Sorting	.15	.20
Test Battery (Unger-Treat)	.27	.09
Senior Classification	.06	.23
Senior Verifying	.10	.19
Minnesota Clerical, Numbers	.26	.22
Minnesota Clerical, Names	.36	.08
Finger Dexterity	.20	.27
Tweezer Dexterity	.07	.46
Minnesota Manual Dexterity	.08	.31
Minnesota Spatial Relations (time)	.24	.28
Minnesota Spatial Relations (errors)	.39	.14
Tracing (time)	−.31	.45
Tracing (errors)	.16	.06
Weaving	.27	.21
Minnesota Paper Form Board	.32	.17
Otis S–A (I.Q.)	.17	.11

which would yield the maximum prediction. A similar selection of a speed test battery was made for the speed criterion.

Test Batteries

The application of Toops' multiple-ratio technique showed that for the quality criterion, a *quality test battery* including the Minnesota Vocational Test for Clerical Workers (Name Comparison), the Tracing Test (time score), the Weaving Test, the Paper Folding Test, the Minnesota Spatial Relations (time score), and the Minnesota Paper Form Boards A and B would give the best prediction. The partial regression equation using these variables was determined; the β coefficients, the gross score combining weights, and the constant can be

found in Table 6. The multiple R for the *quality test battery* with the quality criterion was found to be .57.[16] If a predicted score of 4.80 had been set as the minimum acceptable for admittance to training, 25 individuals would have a criterion score of 4.80 or better, and 8 individuals thus selected would have a criterion score below 4.80.

TABLE 6

Validity Coefficients and Combining Weights for the Tests of the Quality Test Battery

TEST	CORRELATION WITH QUALITY CRITERION	β COEFFICIENT	GROSS SCORE WEIGHT
Minnesota Clerical, Names	.36	.331	.0090
Tracing (time)	−.31	.274	.0015
Weaving Test	.27	−.252	−.0013
Paper Folding	.30	.171	.0120
Minnesota Spatial (time)	.24	.229	.0006
Minn. Paper Form Board A+B	.32	.180	.0071
		Constant	3.4889

It was found that a *speed test battery* including the Tweezer Dexterity Test, Tracing Test (time score), Finger Dexterity, Manual Dexterity, and Minnesota Vocational Test for Clerical Workers (Number Comparison) would give the best prediction for the Speed Criterion, a prediction represented by the multiple R of .64. Table 7 gives the gross score combining weights which were used in obtaining the predicted speed score for each subject. If a predicted speed score of 5.19 or below were chosen, 36 subjects would have been selected for training. Of this number, 27 had satisfactory criterion scores and 9 had criterion scores below the acceptable level.

[16] The use of the multiple correlation technique is usually confined to a larger number of cases than is included in this study. However, it should be pointed out that the 52 cases represent a complete sample of the power operators who finished their training, and this number represents a larger sample than is usually found working on the same type material and machines, and producing the same finished product.

TABLE 7

Validity Coefficients and Combining Weights for the Tests of the Speed Test Battery

TEST	CORRELATION WITH SPEED CRITERION	β COEFFICIENT	GROSS SCORE WEIGHT
Tweezer Dexterity	.46	.348	.0022
Tracing (time)	.45	.412	.0022
Finger Dexterity	.27	.159	.0040
Manual Dexterity	.31	.111	.0029
Minnesota Clerical, Numbers	.22	.090	.0036
		Constant	1.0594

An examination of the scatter-plots showed that a critical score could be selected which, if applied, would have resulted in the selection of better subjects for training. Twenty-one would have been above the critical scores on both quality and speed, and four individuals would have been above on neither. Twelve individuals were above the critical score on the quality battery but were below on the speed battery. The rest of the group, fifteen, would have been selected on the basis of the speed battery but eliminated on the basis of the quality test battery.

GENERAL CONCLUSIONS

It is apparent that the contributions of this study are:

1. The development of reliable and generally objective criteria for *quality* and *speed* of work in garment machine operating, and
2. The development of independent test batteries for predicting *quality* and *speed* of work, respectively.

There seems reason to believe that the use of independent batteries for predicting success on a specific job in terms of quality and speed of performance, respectively, is a novel approach in the study of occupational adjustment. Justification for this approach is to be found both in the practical situation in industry, where speed and quality are given different emphasis by different employers, and in certain psychological studies on the relationship between quality and speed of per-

formance, which show that in many instances it is necessary to treat these as independent variables.

This point of view is emphasized, for example, by Popplereuter, who found distinct differences among apprentices with respect to "speed" and "accuracy" sets as learning proceeded. An analysis of individual work curves by this investigator showed that there are apprentices who work slowly and accurately at first, but who show a distinct trend in the direction of fast and inaccurate work as learning proceeds. On the other hand, there are apprentices who at first appear unsatisfactory, but who ultimately develop into satisfactory workers from the viewpoint of both quality and speed of production. Other types include those with consistently slow work of good quality, workers whose production is both fast and inaccurate, etc.[17]

Although such trends can be observed in work curves, and can be observed in the production of finished work samples, such as the dress used in this investigation, the problems of aptitude measurement and the prediction of vocational success are complicated by the difficulty of clearly differentiating tests with respect to their basic requirements in terms of speed and accuracy. This difficulty has been emphasized by Cox[18] who points out that a universal weakness, where analysis has been attempted, resides in the fact that it seldom penetrates beyond the descriptive level. "A typical example," he writes, "is the common classification of motor tests into 'speed' tests and 'accuracy' tests. On the contrary they are entirely overlapping, for all 'speed' tests need some degree of accuracy, and all the so-called 'accuracy' tests must be carried out with reference to time." That such overlapping exists in the case of individual tests cannot be denied. However, the empirical selection of *speed batteries* and *quality batteries*, as has been attempted in this study, overcomes to some extent the difficul-

[17] Popplereuter, W., "Die Arbeitskurve in der Diagnostik von Arbeitstypen," *Psychot. Z.*, 3 (1928), pp. 35–51.

[18] Cox, J. W., *Manual Skill*, Cambridge University Press, 1934, p. 10.

ties of such an analysis and yet remains within the bounds of good practice as dictated by our knowledge concerning the diversity of observed relationships between *speed* and *accuracy* in different activities.

In general, there seems good reason for agreeing with Cox that there is need for further analysis of the entire problem of the relationship between *speed* and *accuracy* in motor tasks. However, insofar as the problem of occupational adjustment is concerned, it must still be recalled that in the practical situation some employers place an emphasis on speed and others place an emphasis on quality of production. With this in mind, and until we are able to lay down general laws concerning the exact relationship of speed and accuracy, there seems to be real justification, especially in occupational adjustment, for using independent criteria and test batteries for speed and quality.

ONE HUNDRED SUPERIOR MEN

F. L. WELLS, RUTH WILLIAMS AND PHILIP FOWLER[1]
Boston, Mass.

SUMMARY

A GROUP of men with higher business incomes was compared with a group having lower business incomes, by various test procedures, including general intelligence, vocabulary, attitude and others. Such an inquiry is directed at psychogenic factors in the economic success, but not at cultural ones. The high and low salaried groups are practically unseparated by the conventional methods of testing formal intellectual functions. Verbal attitude tests show trends toward separation, along expected lines; immediate environmental factors seem in the main responsible for these. The greater separation appears in procedures giving greater qualitative freedom to the responses than was the case elsewhere, and whose determinants are probably of earlier and more stable origin.

Where the upper brackets of mental ability are being studied, quantitative factors must be relatively discounted in favor of qualitative, and fixed responses in favor of free responses.

[1] The psychometric data were furnished by the business organization with which Williams and Fowler were associated. From the organization's files, Williams collated the basic data for group H, and made some of the initial tabulations. Fowler collated the basic data for group L. For the remainder of the work Wells is responsible, with student assistance from Miss Ruth Felker, Radcliffe '36, and Miss Mary Pomeroy, Radcliffe '37.

OBSERVERS AND TECHNIQUES

Superior, in the terms of the culture in which they live. They are men of business residing in New England, who during the "depression" years from 1929 to 1935 maintained their annual *earned* incomes at a level of $5,000 or more. The superiority denoted is thus limited to economic adjustment in a particular milieu, and is without ethical connotations.

Comparison is attempted with a like sized and aged group, selected from the same general milieu, and similar with respect to cultural opportunities, who had been unable to maintain in business the economic sufficiency denoted by the "superior" material. The higher salary group will be termed group H, the lower salary, group L.[2]

No person was admitted to group L whose earnings since 1929 had averaged over $1,800 yearly. No person in either group was examined for the purpose of the present study.

The tests considered have survived five years of use on account of the light they throw on the personality organization of individuals as here studied. Their function has been to serve as a guide to counseling those persons with respect to business effectiveness, a broader cultural background, understanding of self and others, etc. The tests concerned are such as *in the hands of their users* proved effective for those purposes.

The tests concerned are briefly as follows, the names by which they will be known being given in *italics:*

1. The *alpha* general intelligence test, as revised by Wells ('30) with consideration of total score, amount of error, performance in subtests. It is, of course, taken as an individual test. The scores in a few of the later cases are based on the subtests included in the abridgment published by Wells and

[2] Group H was classified as to annual earnings of (a) $5,000–$8,000; (b) $8,000–$15,000; (c) $15,000 and over. The distribution in these respects is as follows:

Salary class	a	b	c
Number of cases	51	41	9

Atwell (*Personnel Journal*, 1933, 12, 160–163) *i.e.*, subtests 2, 4, 7, 8. In these cases it was, however, necessary to continue using subtest 5, because of usefulness for the general purposes in hand, as a cue to effectiveness with intuitive thinking.

2. The *Inglis* vocabulary test. The reason that this was given instead of the O'Connor revision thereof, is that other considerations made an alphabetical order necessary. For purposes such as the present there is also no gain in arranging the words in order of difficulty, as the entire series must be gone through in any case. Scoring is in terms of words correctly and wrongly marked; general meaning as in O'Connor (*Atl. Mo.*, 1934, 153, 160–166).

3. The *Bernreuter* Personality Inventory. While fully recognizing the force and pertinence of such criticisms as those of Landis (*Journal of Nervous and Mental Diseases*, 1936, 83, 125–134), a procedure of this sort has a place in a group of techniques where it is not taken at face value for the personality involved, but is considered as a verbal pattern, and in the light of other information. For the purposes of this presentation the scores, as in the tests generally, have to be considered as such, and do not give a picture of its rôle in dealing with the individual personalities, in the context of other data.

4. Reversed *Clock*. A procedure based on (1916) Stanford-Binet 14–6, and derived immediately from the work described in Wells and Hylan (*American Journal of Orthopsychiatry*, 1935, 5, 286–301). Instruction to the examinee are as follows:

> A watch or clock must not be looked at during this test. All figuring in this test must be done mentally. Get the time to the nearest minute. Figure in your mind exactly where the hands of the clock are at 6:22, twenty-two minutes past six. Suppose now the hands should change places, the small hand going where the large hand was, and the large hand going where the small hand was. What time would it read then?
>
> You will be given a set of five of these problems. In each case reverse the hands (in your mind) and place your answer beside the question.

The test is scored by (a) the total time consumed in the process, (b) a point score of accuracy ranging from 0 to 20.

It serves in estimating (1) powers of visual imagery, (2) tendency to impulsive judgment. Clinically, it is considered especially in connection with performance in a paper folding test.

5. A form of *overstatement* (and understatement) test. The score here concerned is in terms of a ratio; scores over 100 denoting overstatement, less than 100, understatement. The overstatement technique has for some years been recognized as a useful one, and the present application seems to be specially effective. Clinically, it serves as a cue to (1) conscious estimate of intellectual range, (2) decisiveness or hesitation, (3) desire to impress.

6. A *syllogism* test based on work of Thurstone's, and derived immediately from Wells and Hylan ('35). Instructions to the examinee are,

Consider this reasoning;
Jones is taller than Smith; Brown is taller than Jones; *Therefore,* Brown is taller than Smith.—This reasoning you can easily see is correct. Brown is the tallest of the three.
Now considering this reasoning;
Smith is taller than Brown; Jones is taller than Smith; *Therefore,* Brown is taller than Jones.—This reasoning is incorrect, for Brown is here the shortest of the three.
You will be shown ten sets of reasoning like these. Some of them are correct and some are incorrect. Take them in order. For each one that is correct, write + opposite its number on the small paper given you. For each one that is incorrect, write – opposite its number. Do any figuring on the *small* paper that you need to.

Score is in terms of (a) total time consumed, (b) erroneous judgments. Serves mainly as a cue to ability in manipulating abstractions; clinically to be taken in connection with alpha subtest 7.

7. A test similar to the *paper-cutting* of the original Stanford-Binet 18–2, except that the paper (letter size sheet) is folded once more (*cf.* revised Stanford-Binet). Old Stanford-Binet 18–2 is given first, and irrespective of (Stanford) success or failure therein, the additional fold is tried

with a fresh sheet of paper. Serves in estimating visual imagery. Scoring is by letter grades, as follows: A = all holes and folds correct; A– = Stanford-Binet form correct, but creases only correct with additional fold; B = Stanford 18–2 correct but entire failure with additional fold; C = creases correct, holes incorrect in Stanford-Binet form.

8. The *Hepner* "Executive Reaction" test (Hepner, H. W., "Psychology in Modern Business," N. Y., 1930, xv + 723, pp. 50–54, 699–700); a multiple choice questionnaire on feelings and attitudes, essentially aggressiveness and unaggressiveness, scored on a point basis. Serves mainly in estimating the individual's conscious attitudes towards business activity.

9. An *imagination* test patterned after a procedure of O'Connor's, but of earlier origin.[3] One is asked about happenings or conduct in an imaginary contingency. O'Connor's formula was, substantially, What would happen (what would you do, what do you think others ought to do) if (every electrical conductor suddenly ceased to function)? The examinee is given a specific time within which to write his answer. The time here used is five minutes.

Scoring is by the examiner's rating, in letter grades A–B–C–D. Serves in estimating powers of analyzing an immediate and complicated life situation, subject to the verbal expression of such analysis.

10. A test whose name for record purposes has been *Information,* but which actually consists in verbal definitions of more or less complicated terms such as *dietetics, gneiss, hedonism, semaphore.* More or less "information" is disclosed in the character of the definitions, but the performance is rather a function of verbal freedom and thinking habits. Scored by letter grades, A–B–C–D.

[3] *Cf.* Bronner, Healy, Lowe, Shimberg: A Manual of Individual Mental Tests and Testing. Boston, 1917, (pp. 32–33, *et seq.*). In the present instance the emphasis is laid clearly on what ought to be *done* rather on what would *happen.*

General Comparisons

Various psychometric functions are presumed to be affected by age; and one limitation of the present data is that the age distribution affords at scarcely any point as large a number of cases as one would wish for tracing the effect of age on the functions involved. The age distribution of the two groups is as follows:

Age	25–29	30–34	35–39	40–44	45–49	50–54	55–59	60 up
H	4	5	36	28	16	8	2	2
L	25	26	29	12	6	3	0	0

As will appear below, the test functions concerned do not here exhibit particular relations to age. It has seemed best however, to represent the general comparison of groups H and L through the medium of the age class 35–39 years. This is the age class with the most cases available for both H and L. The H group cases number 36 practically throughout. The L group were not tested in so regular a fashion, and the number of cases varies, averaging in this age class about 22. The averages and standard deviations for the scores in the quantitative test functions are as follows:

	H	*L*	H	*L*	H	*L*	H	*L*	H	*L*	H	*L*
	alpha		alpha errors		Inglis		Inglis errors		Ovstm.		Hepner	
Av.	156	*146*	9.1	*8.5*	131	*131**	14.9	*14.5*	104	*105*	33	*13*
S.D.	21	*27*	5.1	*7.6*	10.4	*14.0*	8.5	*8.1*	6.2	*14.1*	33	*26*

	Bernreuter											
	staby.		ssfy.		domc.		sylg. T		H sylg.err.	*L*	H ClockT.	*L*	H ClockAcc.	*L*
Av.	34	*49*	49	*50*	64	*57*	327	*369*	1.3	*1.0*	218	*252*	9.4	*10.0*
S.D.	29	*26*	27	*27*	26	*25*	144	*161*	1.6	*1.3*	96	*193*	4.8	*3.6*

* On account of its distortion of the picture, especially with respect to errors, there is excluded from the Inglis group L functions, one case who marked all the items, having 67 right and 83 wrong.

It may be observed that there is nowhere a difference of the magnitude known as "statistically significant," although the better average scores are generally found in the H group. In

the formal intellectual processes the differences are negligible as concerns individual cases. A nearer approach to significant difference in these functions is found in the scores of Bernreuter stability and dominance, and the Hepner scores. In certain functions the standard deviations are notably higher in the L group than in the H. In addition to the higher variability thus denoted for the L group, this raises the question of a greater "scattering" of the performance; that is, wider gaps between the best and poorest functions in a single individual, than obtain in the H group (cf. below, under "Psychometric Balance").

From the aspect of the total cases, there is probably some small difference for the alpha averages of groups H and L, but nothing of individual significance. The alpha errors show less difference than might be expected, and do not seem consistent with a hitherto observed trend for intelligence test accuracy to be associated with generally better personality integrations. (It is perhaps a gratuitous assumption that the mental hygiene of group H is better than that of group L). Although the differences are "statistically" negligible, it seems worth noting that the accuracy scores in both *Syllogism* and *Clock* tests, with their standard deviations, are in favor of the L group.

These results are in conformity with the accumulating evidence that the quantitation of formal intellectual processes is of very limited significance for the integrations of the personality as a whole.

Special Comparisons

In the Hepner scores for the age class 35–39 already cited, the difference of groups H and L was of little statistical significance, but examination of the total material lends itself to a more favorable view. The distribution of scores is given on the following page.

There is observed some tendency for the Hepner scores to increase with age. Also to be noted, is the high position of

	to −40		−39/−20		−19/−1		0/19		20/39		40/59		60/79		80 up	
	H	L	H	L	H	L	H	L	H	L	H	L	H	L	H	L
25–29		4		6		5	2	8	1	8	1			1		
30–34		8		7		6		5		4	1	1	4	1		
35–39	1	1	2	8	1	4	8	10	12	7	5	3	3	1	4	
40–44				3	3		7	5	2	3	7		6		3	1
45–49			1	3		1	4	3	2		5		4			
50–54			1			1	3		3	3	1					
55–59									1		1					
60–64											1				1	
Totals	1	7	4	19	4	17	24	24	21	23	22	4	17	3	8	1

the few H group cases under 35, and the comparative rarity with which these levels are attained by the more numerous L group. The main point of interest concerns the rôle of psychogenesis, and the issues can be at least indicated, though the solution is a matter for individual study. It is suggested, however, that the (*verbal!*) dominant attitudes characterizing the age classes 25–34 of group H must have been pretty well laid down before their responsible business careers began, must therefore represent a transfer from dominance in other types of situation. From age 35 on this dominance pattern appears rather a function of business prestige, becoming ingrained as age, and perhaps business status, advances.

The L group lends itself less distinctly to psychogenic understanding. The absence of the pattern in the early years is intelligible enough, but less so even the slight emergence in the later years, particularly in the face of the relative business frustrations under which they lived at the time of the observations. Compensatory mechanisms may be involved, and there may be a temperamental maturation expressing itself in a plus direction of the responses.

The H group, and those of the L group for whom the data are available, are here separated by ages according to record of college graduation or its absence:

Ages	Less than Graduation		Graduation	
	H	*L*	*H*	*L*
25–29	2	*6*	2	*18*
30–34	1	*9*	4	*15*
35–39	15	*15*	21	*7*
40–44	14	*7*	14	*4*
45–49	8	*2*	8	*4*
50–54	3	*1*	5	*2*
55–59	1		1	
60–64			2	
Totals	44	*40*	57	*50*

It will be seen that the groups are practically equal in this respect. It is not possible to infer direct influence of college graduation one way or the other. It is possible, however, that non-college graduates are likely to achieve H status more on their "own merits" than are college graduates. In so far as this is true, the rôle of college graduation may even be slightly negative for H status.

Another natural supposition would be, that non-college graduates might have relative advantages in earlier years, but that the graduates would later wipe this out. There is no evidence of this in the present data. There is also the possibility that personal characteristics favorable to H status react negatively to college training of the liberal arts type.

Of the three tests scored by letter grades, *information* is lacking in data for group L; the group H data are unaffected by age. The paper tearing tests afford no basis of distinction between group H and L. 33 per cent of group H, 42 per cent of group L do better than the passing grade of 1916 Stanford-Binet 18–2; but this is again chargeable if not to chance, to the overweighting of group L at the younger ages.

The nearest approach to a separation of groups H and L is found in the scores of the *imagination* test. These scores, it may be remembered, represent the richness and reasonableness of responses to a question of what would happen or what should be done in a hypothetical emergency or other situation. Assurance may be given here that there is negligible halation from experimenter's knowledge of the examinees' business status, and the four ratings, A–B–C–D, are themselves endowed with a fair degree of objective distinction. D represents practical failure, only a line or so of writing being produced, and this of a vague character. Responses of A grade might run to 150 words or more per question. No attempt is made to present the ratings quantitatively, though a reasonable trial could be made with the equivalents, D = 1; C = 4; B = 16; A = 64. Distribution by ages is as follows for the 101 H cases, and L cases numbering 66:

	A		B		C		D	
	H	*L*	H	*L*	H	*L*	H	*L*
25–29	2		2	*10*		*8*		*3*
30–34	1	*2*	3	*6*	1	*10*		*5*
35–39	4		21	*7*	10	*2*	1	*5*
40–44	5		17	*1*	6	*3*		*2*
45–49	1		6	*1*	9			
50–54			5	*1*	3			
55–59			1		1			
60–64	1		1					
Totals	14	*2*	56	*26*	30	*23*	1	*15*

Scores of A are made by 3 per cent of L cases against 14 per cent of H. D scores are made by 23 per cent of L cases against 1 per cent of H. In the B and C grades the differences are hardly significant; but when one observes an A response there is reasonable probability that H group qualities are involved, and when one observes a D response there is strong probability that the qualities are nearer those of group L.

There is some possibility that the older examinees score less well, but age trends are rather conspicuous by their absence. It would "make sense" for the responses to be more vivid and creative in the younger individuals, and to be broken down by "reality" in advancing age, this process being reinforced in group L by the comparative frustration to which they have been subject. But the danger in rationalizing trends so indistinct is considerable; growing experience in group H could well lead to richer responses, and the frustration of group L be a stimulus to imaginative compensation. As elsewhere in this material, the question calls for study at clinical rather than statistical levels.

It will be noted that this procedure is distinguished from the previous ones in the degree to which the observer organizes his own responses. Tests like *alpha* or *Inglis* are dual or multiple choice. Tests like the *Clock*, while the response is not

given ready-made, still restrict the response within such narrow categories as to be multiple choice in principle. In the present *imagination* procedure, the response range is limited only by the observer's associative powers, as in procedures of the Rorschach type; and is naturally more adapted to disclosing differences in those initiative and originating aspects of the personality that make for presence in group H.

It was pointed out at the beginning that the distinction of groups H and L had no ethical significance. It is relevant to comment briefly on its psychological significance. The conventional view would be that group H rose higher because their inherent personal traits were such as better fitted them for success in a competitive pecuniary culture. This is a view that fits but partially the operation of a relatively crystallized social structure, in which status somewhat inhibits competitive selection. It is not impossible thereunder for very moderate abilities to achieve and maintain, under protection, the economic qualifications for group H; and for higher qualities in group L to be frustrated by circumstances, "the resistance of vested things."

One should not, therefore, expect on psychometric grounds clear separation of groups selected as are the present H and L, for the reason that they are not and cannot be selected exclusively on the basis of traits distinctly personal in character. Hereditary and environmental accidents of all sorts make contributions not directly accessible to psychometric observation. Whether more stress on methods of mental measurement would permit more socially benign career selections than do the usual mixtures of static and opportunistic cultural factors, is impertinent to this discussion; but there is no reason to suppose that such selections would be similar.

INTERRELATIONS OF TEST SCORES

(From this point on, data refer to group H only.) In any discussion of correlations in the above scores, it should be kept in mind that one is dealing with an upper stratum of the intel-

lectual functions. The median intelligence test score is above the 95th percentile for the population at large. Such correlations as would be evident for a sampling more representative of the general population, are therefore largely wiped out. For the following pairs of variables examined, the relationship appearing is for present purposes insignificant.

Variables	Remarks
Alpha × Education	The median scores of graduates and non-graduates differ by only 6 alpha points, though the non-graduates have five times as many cases scoring below alpha 140.
Alpha × Hepner	Of the 8 cases with Hepner ratings over 80, 1 was among the 39 cases below 150 alpha.
Alpha × Imagination	
Alpha × Salary	No member of the class earning over $8,000, 6 members of the $5,000–$8,000 class, scored below 120 alpha.
Inglis × Occupation	With reference to the proportion of sales and administrative work.
Inglis × Hepner	A good example of rectangular scatter.
Inglis × Syllogism Time	Excessively long Syllogism Time occurs with Inglis scores below 85; above Inglis 95, Syllogism Time scores vary with substantial independence.
Inglis × Imagination	Of the 14 A grades in imagination, none has Inglis below 115.
Inglis × Salary	Of the 9 cases in the $15,000+ salary class, two are below 100 Inglis, three above 145 Inglis.
Age × Alpha Errors	
Age × Inglis Errors	
Age × Clock Speed	
Age × Imagination	
Age × Salary	No case in the $15,000+ class was under 35; no case in the $8,000–$15,000 class was under 30.
Age × Overstatement	
Occupation × Hepner	
Occupation × Alpha Errors	
Occupation × Imagination	

Occupation × Overstatement	Extreme overstatement more associated with limited formal education.
Education × Overstatement	
Education × Salary	I.e., substantially no difference in salary grouping according to fact of college graduation.
Overstatement × Bernreuter Self-sufficiency	
Overstatement × Bernreuter Dominance	
Overstatement × Hepner	

In other cases there is some indication of common factors, or a relationship is already known to exist, as between vocabulary and "intelligence" scores.

Variables	Remarks
Alpha × Inglis	The known relationship is sustained in the lower ranges of scores, but breaks down in the higher. In the bulk of cases, which are above 130 alpha, no relationship whatever appears. This is important, because when such a person's vocabulary is too high for his alpha, or vice versa, it has implications for training.
Alpha × Age	The four cases under 30 all have alphas over 170; in general the intelligence score stands up better than in findings such as Miles'. The view is confirmed that intellectually stimulating occupations largely arrest the decline of score with age.
Alpha × Syllogism Time	The central tendencies of the successive steps progress regularly.
Inglis × Overstatement	Marked negative relationship shown in the scattergram is partly spurious, but tendency to higher overstatement scores in the lower vocabularies is genuine.
Inglis × Education	Number of non-graduates with Inglis scores under 120, 13; of graduates, 5. Non-graduates with Inglis scores above

	140, 4; of graduates, 16. (Total, 44 non-graduates, 57 graduates).
Age × Clock Accuracy	Tendency for the higher ages to have *slightly* lower scores.
Education × Syllogism Time	The graduates have a median 14 per-cent faster, but the difference is much less clear-cut than with the alpha scores.
Overstatement × Alpha Errors	The number of alpha errors made by those understating is about two thirds the number in those overstating.
Overstatement × Inglis Errors	The number of Inglis errors made by those understating is about one fourth the number in those overstating.
Overstatement × Clock Accuracy	The median accuracy score of those understanding is about one and one-half times that in those overstating; the extremes differ as 6: 1.

PSYCHOMETRIC BALANCE

The profile of the functions studied has been indicated; the complete alpha median approximates 155, etc. There presents itself the problem of variations about the norms. The more closely an individual profile approximates the norms, the more "balanced" these functions are, in a statistical sense. Does this also denote better balance in an adjustmental sense, in the common meaning of "well-balanced" personality? If, for example, one rates very high in *alpha* and *vocabulary*, and extremely low in *imagination* and reversed *clock*, with similar spread between other psychometric functions, does this denote an inferior balance in an adjustmental sense?

A number of years ago, one of the writers participated in a frontal attack on this problem, from a somewhat different angle (Wells and Sturges, *American Journal of Insanity*, 1918, 75, 81–119), since psychometrics as we know it now did not exist. Close comparison was made in respect to choice reactions in functions more specialized than here, between persons of superior adjustments, and others suffering from serious psychotic disorders. At this level of response no real differ-

ence appeared, and it is of course commonplace that sufferers from various psychoses may be excellent pianists, bridge players, or performers with intelligence tests. But this is not saying that their performance in these functions would maintain a *Gestalt* consistent with normality; by definition it would not do so, if one understood the critical measures to apply. In a narrow way, something has been learned about the meaning of "scatter" in tests of the Binet type, and it has been looked on as associated with relatively disturbed functioning. In the present material, such "scatter" is more precisely determinable. What relation can be traced between it, and the quality of life adjustment among members of this group?

The conventional measures for this purpose are S.D. units, and a profile plotted in these terms is a direct statement of the degree of balance between psychometric functions. The mean degree of this deviation offers a single figure to represent it, where this is desired. In terms of such a figure, the amount of this scatter distributes itself as follows:

Average of an observer's sigma deviations in the several tests	.4	.5	.6	.7	.8	.9	1.0	1.1	1.2
Number of cases	3	6	13	21	15	17	14	5	1

(*Continued*)

Average of an observer's sigma deviations in the several tests	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.
Number of cases	1	1			1	1	1	1

Does the relative psychometric imbalance shown by the figures above 1.2 reflect an imbalance in more general features of personality adjustment for the individuals concerned? When the nine below .6 are compared with the seven who ranked above 1.2, the available comparisons disclose no direct relation between psychometric and personality balance; there is indeed rather more neuroticism among the less scattering group.

The following table gives scores approximating certain mul-

tiples of S.D. for group H, age class 35–39, in alpha, Inglis and Hepner scores. They are based on the table of averages cited for this age class, and give standards for a sort of *corps d'elite* in the functions concerned.

Sigma Units	Test Scores		
	Alpha	Inglis	Hepner
−2.0	114	111	−33
−1.5	124	116	−17
−1.0	135	121	0
− .5	146	126	17
+ .5	166	136	50
+1.0	177	141	67
+1.5	187	146	83
+2.0	198	(151)	99

There is little indication, however, that the extension of this procedure to the other tests would distinguish a *corps d'elite* in business efficiency. This may be exemplified by the four cases under thirty who had achieved group H status. If one compares their test status with the total group averages, one finds four functions, out of sixteen, in which all were better than this average (alpha, syllogism time, syllogism error, clock time); three cases in which but one was better than average (Inglis, executive reaction, overstatement ratio). While it is most improbable that these particular patterns are the result of "chance," they are not far from what chance might produce; and minimal significance is indicated for the general *amount* of deviation from the group averages.

It may be further remarked with regard to these cases, that there is evidence neither of favoritism in their early advancement, nor of failure of business ability in sustaining its responsibilities. They represent some contrasts in personality types, which may be sketched as follows:

(W) A man of much physical energy and particularly analytic mind; in younger days very shy, history of greatly

"introverted" pattern in a word association test. Successfully disciplined this business handicap, perhaps to a certain overcompensation. Now at 33 years is a major industrial executive.

(X) Highly conscientious devotion to job; till recently much fatigued by day's work, apparently through resulting somatic tensions. Abilities attracted attention of superiors, who gave "horizontal" promotions as preparation for considerable advancement. Rate of progress has been accelerating. Attitude towards work now less tense, remission of fatigue effects lengthens average day by 1½ hours. Very amiable personality.

(Y) Circumstances leading to group H status were more fortuitous here than in the others, but the duties were satisfactorily performed. Personal dislike of the duties led to voluntary cessation of connection.

(Z) Emotionally stable and aggressive hypomanic. Good opinion of self, ambitious for position and power; trend also to hero-worship. Rapid rise on merits, earning $6,000 a year at 28 years; looked on as highly promising man of business. Subsequent death, from accidental injuries.

TEST PATTERN OF MENTAL DEFECTIVES SKILLED IN WEAVING

ROY HAMLIN

Letchworth Village, Thiells, New York

AND

THEODORA M. ABEL

Director of Research, Trade Extension Classes, New York City

THE purpose of this investigation was to find whether successful and unsuccessful weavers could be differentiated on the basis of test measurement. A differentiation of this kind we felt would throw light on the kind of special ability or mode of functioning necessary for weaving on hand looms.

Weaving is one of the occupations provided for mentally deficient girls in the industrial school of Letchworth Village, a New York State institution.

A previous paper reported an investigation of another occupation, pillow-lacemaking. A small group of successful lacemakers was compared with another small group of failures at lacemaking. Both groups had expressed a desire to learn to make lace and had been given an equal chance to succeed. On the Stanford-Binet and on a battery of non-language and performance tests, no reliable difference was found between the successes and failures.[1] But, on a series of simple non-standardized tasks (i.e., card sorting, tapping, marking short lines in squares of graph paper), the two groups were definitely differentiated. This difference, however, was not brought out by scores or averages of scores on any task or group of tasks but by what was termed for convenience a "pattern score."[2]

[1] "An Analysis of Capacities for Performance of Mental Defectives Skilled in Lacemaking. I. A Study of Standardized Test Results," G. C. Taft and E. F. Kinder, *Journal of Applied Psychology*, XX, 5, 567–575, 1936.

[2] "An Analysis of Capacities for Performance of Mental Defectives

The procedure in obtaining these patterns was as follows: The subjects worked at an initial task to which they became somewhat adapted, then at a second or new task that was very similar to the first task but enough different so as to require some further learning, and finally they returned to the initial or basic task. From these tasks was derived the performance pattern, that is, the relationship of scores on the new task to scores on the basic or more familiar task. The successful lacemakers did relatively better on the return to the initial task than on the new task. This marked difference seemed to be due to some fundamental difference in ability regardless of the intelligence level of the lacemakers and failures. That is, those who had sufficient ability to work ultimately very quickly but who actually worked very slowly while learning were the ones who succeeded in lacemaking.

The present problem was to study another occupational group, the loom weavers, and to see whether a test pattern that would differentiate the successes from the failures could be found. The plan was to employ standardized tests rather than unstandardized tasks, and to find among the tests some pattern of performance that might characterize success and failure.

The subjects were mentally deficient girls in the industrial school who had been given an extended period of instruction. The successful group consisted of twenty subjects with chronological ages from ten to thirty-two years. This group had succeeded in mastering the performance of weaving and could work out simple patterns. The group that failed consisted of ten subjects with chronological ages from fourteen to thirty-nine years. This latter group had been tried on the looms as long as had the successful group but had failed to grasp the processes involved in working even the simplest loom. Their failure was complete. All intermediate cases that were classified as partial failures were excluded from this experiment. On Binet mental age scores, the two groups, successes and failures were fairly equivalent.

Skilled in Lacemaking. II. A Study of Performance in Selected Non-Standardized Tasks," T. M. Abel and R. Hamlin, *Journal of Applied Psychology*, XXII, 2, 175–185, April 1938.

The next step was to give a battery of seventeen tests to the subjects. These tests were: Manikin feature-profile; digits; Ferguson Formboards; Goodenough Drawing; Gwynn Triangle; Substitution; Vocabulary (from the 1937 Stanford-Binet); Healy A; Seguin Formboard; Knox Cubes; Diagonal; Kohs Blocks; Healy P. C. I and II; Glueck Ship; Army designs; Comprehension (1916 Binet). In addition, the subjects were given one of the unstandardized tasks that had been used in the previous study of the lacemakers, the tasks that differentiated the successes from the failures most effectively (marking tasks).[3] But on these tasks no differences were revealed between the successful weavers and the failures.

Among the seventeen standardized tests, however, there were four in which the scores of the weavers were superior to those of the failures, without obtaining a pattern score. These four tests were the Seguin, Ferguson, Knox Cubes and Healy A. The results were as follows:

	NO.	AVERAGE M.A.	SIGMA	RANGE
Successful weavers	20	8.9	.3	6.1–12.7
Failures	10	6.2	.3	5.1– 7.9
Difference		2.7	.4	

This difference is 6.75 times as great as its sigma, indicating a high degree of reliability, though the groups are small. In spite of this marked difference, however, there was still overlapping in scores between the two groups involving ten of the thirty subjects, or one third.

A study of these overlapping cases suggested a test pattern as an even more effective measure of weaving ability. Two outstanding examples will illustrate the findings. The subject with the highest Binet mental age of any of the thirty had failed at weaving, but presumably because of her high general intelligence, had made the top score among the failures on the four critical tests. On the other hand, the subject with the lowest Binet mental age of any of the thirty had succeeded at

[3] *Op. cit.*, p. 178.

weaving. But apparently because of her low general intelligence, she made the lowest score of any successful weaver on these four critical tests.

It seemed that the four critical tests (Knox Cubes, Seguin, Ferguson, Healy A) measured partly the abilities or modes of functioning required in weaving, but also partly other abilities associated with general intelligence.

The next step was to select tests of intelligence on which the successes and failures scored about equally. Among the other tests, three met these requirements the most adequately. These three non-differential tests were the Vocabulary (from the new Binet), the Goodenough and the Substitution tests. For the successful weavers the average mental age on these three tests was $7.5 \pm .3$, and for the failures this average was $7.2 \pm .5$. The difference in average score is only $.3 \pm .6$. This difference is only .5 times as great as its sigma, indicating a negligible difference between the two groups.

Instead of a subtraction technique, as with the scores of the lacemakers and failures in the previous study, a ratio was used as the pattern score. This ratio was $\frac{X}{Y}$ where X equals the combined mental age scores on the four differentiating or critical tests, and Y equals the combined mental age scores on the three non-differentiating tests. The results were as follows:

	NO.	AVERAGE RATIO SCORE	SIGMA	RANGE
Successful weavers	20	120.6	3.8	102–173
Failures	10	87.6	2.4	73–100
Difference		33.0	4.5	

The difference between these average scores is 33 ± 4.5, a difference which is 7.33 times the sigma. Furthermore, this pattern score distinguished between the two groups so effectively, that every one of the twenty weavers surpassed every one of the ten failures. There was no overlapping between the two groups.

Apparently, the successful weavers are individuals who have some particular ability or abilities that are not clearly mea-

sured by individual tests. Some of these tests, however, do measure apparently the special ability of the successful weavers plus other abilities they have in common with the failures at weaving.

Just what these particular abilities are that make for success in weaving and that are revealed by means of pattern scores of standardized tests is not easy to determine without a good deal more research and careful study of a larger number of cases. It may be suggested, however, that this ability is related in some way to the perception and understanding of simple spatial relationships. Certainly, solving the formboard problems and carrying out the activity of the Knox cube sequences requires ability in observing and handling spatial relationships, in addition to general intellectual ability. Also, being able to weave on hand looms requires perception of and manipulating spatial relationships.

In this and the previous study, a careful and extended study of scores on many tests was required before a satisfactory pattern was worked out. In general practice, the possibilities of such careful inspection of test scores may be neglected for this reason: the investigator is inclined to pass over fairly readily the scores on tests that seem to distinguish not at all, or only very slightly, between two groups of subjects. He is also somewhat inclined to pass over tests that show a slight difference in favor of a group of "poor" subjects, i.e., failures at a given occupation. It is not readily apparent how low general intelligence helps indicate probable success at weaving, or how a poor showing on an unfamiliar speed test helps indicate probable success at lacemaking.

It is possible that much can be done with test patterns in measuring occupational aptitudes or abilities that has not been done in the past because of this neglect of tests that do not have any readily apparent importance. The present study may serve as a beginning in this direction in differentiating successes and failures at weaving, and in indicating possibilities for developing pattern scores as measures of occupational aptitude.

A SIMPLIFIED SCORING METHOD FOR THE BERNREUTER PERSONALITY INVENTORY

GEORGE K. BENNETT

The Psychological Corporation

THE great frequency of use of the Bernreuter *Personality Inventory*[1] is evidence of the felt need for a measure of this sort. This is especially true when the high cost of scoring this questionnaire is considered. The long time required for scoring this blank is due to three factors: 1. It is scored for several scales (4 or 6); 2. Scoring on each scale requires the algebraic summation of 125 individual weights ranging from −7 to +7; 3. A percentile equivalent must be ascertained for the score on each scale.

If the Personality Inventory is scored manually for the four scales established by Bernreuter (neurotic tendency, self-sufficiency, introversion-extroversion, dominance-submission), from fifteen to twenty minutes per blank is usually required. When the two Flanagan factors[2] are included, the scoring time ordinarily ranges between twenty and thirty minutes. There are three methods more or less frequently used to reduce the time and cost of scoring.

The most elaborate of these methods involves the use of Hollerith equipment with special attachments. The usual charge for scoring one Personality Inventory by this method is fifteen cents plus return postage and there is the additional disadvantage of time lost in mailing. Hand scoring is facilitated by the use of two Veeder counters. This method expedites scoring appreciably, but requires the purchase of special

[1] Bernreuter, Robert G., The Personality Inventory, Stanford University Press, 1935.

[2] Flanagan, John C., Factor Analysis in the Study of Personality, Stanford University Press, 1935.

equipment. A seldom used but feasible method is to score the blanks for the two Flanagan scales only and to estimate the four Bernreuter scores by means of regression equations. Although this procedure will save time, few users appear willing to make the necessary computations.

The International Business Machines Company's new test scoring machine is adaptable to this blank. However, unless revised weights are used, three sets of stencils will be required for each scale and the answer sheet will be passed through the machine 18 times for complete scoring.

It is apparent that if the blank could be scored for only the two Flanagan scales, the weights of these scales reduced to small positive numbers, and the location of percentile equivalents simplified, much time would be saved in scoring. Accordingly, the following steps were taken:

1. A tabulation was made of the range of weights assigned to each item in the two Flanagan Scales.

2. Equivalent small positive weights were assigned according to the original total weights. Items having a weight range of 3 or less in the Flanagan Scales are given no weight; those within the range of 4 to 7, a weight of 1; and those with a range of 8 or over, a weight of 2. Where the third answer had an intermediate weight on the original scale it was given the most suitable new weight.

3. Using these small weights, 115 previously scored Personality Inventories were rescored for the two simplified Flanagan factors. The correlation between the original and simplified Flanagan Scales was .971 for F1–C and .938 for F2–S. The intercorrelation of the rescored scales was – .015.

4. The correlation between each of the rescored scales and the original four Bernreuter scales was also determined.

	B1–N	B2–S	B3–I	B4–D
F1–C, rescored	.955	– .714	.930	– .875
F2–S, rescored	.128	.608	.189	.163

5. New regression equations were constructed to estimate

each of the four Bernreuter scales from the rescored Flanagan factors. The equations follow:

$$B1\text{–}N = 4.120\,(F1\text{–}C) + 1.083\,(F2\text{–}S) - 338.553$$
$$B2\text{–}S = -1.917\,(F1\text{–}C) + 2.831\,(F2\text{–}S) - 69.127$$
$$B3\text{–}I = 2.269\,(F1\text{–}C) + .873\,(F2\text{–}S) - 210.362$$
$$B4\text{–}D = -3.034\,(F1\text{–}C) + .917\,(F2\text{–}S) + 119.462$$

The coefficients of multiple correlation between obtained and estimated scores from the above equations are:

B1–N = .965	B3–I = .951
B2–S = .934	B4–D = .889

6. Since only two independent variables appear in each of the above regression equations, it is possible to construct a two-dimensional table permitting one to determine for any combination of F1–C and F2–S scores the estimated scores for all four Bernreuter scales. If one is concerned with a large number of inventories administered to a group homogeneous as to sex and academic status, *e.g.*, high school boys, the cells of the table may contain, instead of raw scores, the percentile equivalents of these scores for that group.

As a check on the accuracy of estimate of this method, percentiles were obtained for a new group of 100 adult women whose blanks were scored both by the Hollerith method and by the procedure here described. Two sets of percentile equivalents were obtained for the four original scales; the first from the standard Bernreuter percentile tables and the second from a chart similar to that shown in Figure 1. For each scale the correlation between percentiles was determined using adjusted class intervals to normalize the percentile distributions. The table below shows the original multiple correlation coefficients for estimating the raw scores and the obtained coefficients between the two series of percentiles.

Scale	*Multiple R*	*Obtained r*
B1–N	.965	.951
B2–S	.934	.883
B3–I	.951	.934
B4–D	.889	.924

Abbreviated F2-S score	F2-S percentile	Bernreuter Scale	COLLEGE MEN Percentile Equivalent								
97		B1-N	1	1	2	6	9	14	21	28	99
	99	B2-S	99	99	99	99	99	99	99	99	42
		B3-I	1	1	3	6	12	16	22	29	99
95		B4-D	99	99	99	98	97	95	92	86	3
46		B1-N	1	1	1	1	1	1	4	8	96
	2	B2-S	70	66	60	54	49	43	37	32	1
		B3-I	1	1	1	1	1	1	2	3	92
44		B4-D	97	96	92	87	82	77	71	64	1
43		B1-N	1	1	1	1	1	1	4	7	95
	2	B2-S	65	59	54	48	42	37	31	25	.1
		B3-I	1	1	1	1	1	1	1	3	92
41		B4-D	97	95	91	86	81	75	69	62	1
40		B1-N	1	1	1	1	1	1	3	6	95
	1	B2-S	59	53	48	41	36	30	25	20	1
		B3-I	1	1	1	1	1	1	1	2	91
38		B4-D	97	94	90	85	80	74	68	60	1
F1-C %ile			1	1	1	2	6	11	18	25	99
Abbreviated F1-C score			4	8	12	16	20	24	28	32	100
			1	5	9	13	17	21	25	29	97

FIGURE 1. Reproduction of portions of table of percentile equivalents for college men. Each column represents a score interval for the first Flanagan factor and each row, an interval for the second factor. The entire table has 25 columns and 20 rows. Similar tables have been prepared for the remaining five sets of percentiles. (High School Boys, High School Girls, College Women, Adult Men, Adult Women.)

It will be observed that the discrepancies are rather greater than would be expected if the equations were applied to a new group of subjects. However, the multiple R's represent estimation of raw scores, and the obtained r's estimation of percentiles. Both sets of percentiles include errors of approximation and the conversion of the percentiles to normal distributions introduces further error of this sort.

The economy of this method from the standpoint of time is striking. Scoring of a blank by the two abbreviated scales requires less than four minutes. Locating and copying the percentile values consumes less than one minute. Thus there is a total time economy of 75 per cent as compared with ordinary scoring.

Since the multiple correlation coefficients are less than unity, there is some loss in accuracy by this method. However, the standard error of estimate from the regression equations is only slightly greater than the standard error from the unreliability of the original scales.

In view of the fact that the traits which we are attempting to measure are subject to large quotidian fluctuations, at least as measured by the original scales, it is believed that the economy of scoring more than justifies the small loss in precision.

SUPERIOR INTELLIGENCE IN PATIENTS WITH FRÖHLICH'S SYNDROME*

EMMETT L. SCHOTT
Henry Ford Hospital, Detroit, Michigan

IN a group of more than 3000 patients referred to the writer during the past ten years for formal psychological studies, approximately thirty have been girls and boys between the ages of nine and eighteen years whose major diagnosis was some form of glandular dysfunction. Of these thirty patients, seven boys have been diagnosed as typical cases of Fröhlich's syndrome. This condition, ascribed to hypo-function of the anterior lobe of the pituitary gland, was, according to Curschmann (2), originally described by Fröhlich in 1901. In Curschmann's book, published in 1929, the condition was mentioned as comparatively rare. According to Novak (9), writing in 1935, the condition is an "extremely common clinical picture." In brief, the physical signs that characterize these patients are,—heavy girdle of adiposity about the abdomen, hips, and buttocks, large breasts, shoulder pads, and the genital hypoplasia outwardly more apparent in the male.

With one exception, the seven boys with whom this paper is concerned were being studied and treated by the Division of Metabolism in our hospital when they were referred to us for psychological evaluation. The one exception was sent to us from the Division of Pediatrics because of behavior difficulties,

* This paper in preliminary form was presented before the Section on Psychopathology, at the eleventh annual meeting of the Midwestern Psychological Association at Northwestern University, Evanston, Illinois, on April 25, 1936. The study has been made possible through the cooperation of Dr. Thos. J. Heldt, Physician-in-Charge, Division of Neuropsychiatry, and Dr. Daniel P. Foster, Physician-in-Charge, Division of Metabolism, Henry Ford Hospital, Detroit, Michigan. The writer is indebted to them and to Jack H. Schott for suggestions regarding revision of the manuscript.

but later had most of the same studies as the others, and was found to have the same characteristics of adiposogenital dystrophy described above. The ages of our seven patients ranged from eleven to eighteen years. Considering their ages as of their last birthday, two were eleven, one was twelve, three were thirteen and one was eighteen years of age at the time of our studies. Both the average and the median ages were thirteen years. In height they ranged from one and one-fourth to eight inches taller than the norms for their ages and as a group they averaged four and seven-tenths inches taller than these norms. In weight, they ranged from three to ninety-eight pounds heavier than the norms for their ages and as a group they averaged forty-nine pounds heavier than these norms. With one exception the basal metabolic ratings were toward the minus side (minus two to minus eighteen per cent), indicating possibly some degree of hypothyroid function. Other recorded physical and laboratory data including blood pressure readings, blood sugar and phosphorus curves, and x-rays of the skull, showed no abnormal findings that were uniform in tendency.

The few psychological studies of hypopituitarism and other glandular disturbances reported in the literature have emphasized intellectual retardation and the possibility of raising the IQ by glandular therapy. Potter (10), on the basis of his work with mental defectives, warned in 1923 that endocrine therapy should not be regarded as a panacea for mental deficiency; in a later article (11), he suggested that, along with desirable physical changes, appropriate glandular therapy would also "stimulate sluggish intelligence" in certain cases. In 1928 an investigation by Fox (3) indicated that some slight improvement in IQ ratings on a group of defectives was due to glandular therapy and that the relationship between glandular dysfunction and IQ was highest in the hyperpituitary group. At about the same time, Menninger (8) concluded that there was no uniformity of association of any mental picture with any type of pituitary disorder, but he stated that

patients with pre-adolescent hypopituitarism usually showed intellectual retardation. In 1930 Marinus and Kimball (5) found that pituitary treatment did not significantly alter the intelligence of children with pituitary deficiencies, although there was an improvement in school work, ascribed to the stabilizing effect of treatment upon the personality. In 1934 Aitkin and Russell (1) mentioned mental deterioration as a characteristic finding in patients with Simmonds' syndrome, which is classified as a manifestation of pituitary anterior lobe deficiency. In the same year, Rubinstein (12) reported a marked rise in the IQ among other desirable results in a patient treated with the anterior pituitary-like hormone. In 1934 Mateer (6) reported acceleration of IQ's in a group of subnormals being treated with pituitary medications. In 1935 the same author (7) found improvement in reading ability in a number of pupils who had been diagnosed pituitary deficients and who were receiving appropriate therapy for this condition.

In contrast to the majority of these studies, our patients of Fröhlich's type are, as a group, strikingly above the average in intelligence. Each one was given a complete examination with the Stanford-Binet test and a few of them had supplementary studies including the Seguin form board performance test. On the Stanford-Binet test, their IQ's were 102, 117, 119, 132, 139, 144, and 159. The median IQ was 132, which is usually considered as being in the upper one per cent of the general population. These seven boys, with the exception of the one whose IQ was 102, were accelerated educationally from one to two years above the averages for their ages.

It is difficult to account for the discrepancy between the intellectual level of our Frölich's patients and the findings of other investigators. The possibility of a selective factor is suggested by the fact that our private hospital patients are drawn largely from the upper levels of the socio-economic scale. However, in connection with other studies, (13) (14), we have tabulated the IQ's of 850 of our patients (including both children and adults) and found normal distributions with the

percentage of superior individuals about the same as the corresponding figures for the general population. The apparent IQ increases which have been ascribed to pituitary treatment may have been due to improved physical condition and emotional state or possibly the relationship involved was obscured by the absence of adequate control groups in the investigations. For example, Marinus and Kimball (5) found that seventy per cent of their patients under pituitary treatment showed improvement on achievement tests. This might have been interpreted as the result of glandular therapy, had not eighty per cent of a control group, diagnosed as pituitary deficients and not receiving any treatment, also shown improvement. Moreover, a hypothyroid condition often is associated with pituitary deficiency and in clinical practice both thyroid and pituitary medications are frequently given concurrently. Since thyroid function is known to have a definite effect on general development (4), at least in extreme cases such as cretinism, it is possible that some of the mental improvement credited to pituitary treatment was due to thyroid treatment received at the same time.

The number of cases in the present study is small and it may be entirely due to chance that our group of patients with Frölich's syndrome show superior IQ ratings. Such data prove only that intellectual retardation is not necessarily a concomitant of pituitary deficiency. It may be that intelligence and pituitary function are not directly related at all. In any event, generalizations concerning the effect of pituitary function upon intelligence should await a better understanding of the separate pituitary hormones and their relationship to specific functions of the organism.

REFERENCES

1. Aitken, R. S., and Russell, D. S. A case of Simmonds' syndrome. *Lancet*, 2: 802–806, Oct. 13, 1934.
2. Curschmann, H. Endocrine disorders. Oxford University Press, 1929, p. 80.
3. Fox, E. J. An investigation of the effect of glandular therapy on the intelligence quotient. *Ment. Hyg.*, 12: 90–102, Jan., 1928.

4. Gregory, Jennie. A B C of the endocrines. Williams and Wilkins Co., 1935.
5. Marinus, C. J., and Kimball, O. P. Endocrine dysfunction in retarded children and their response to treatment. *Endocrinology*, 14: 309–318, Sept.–Oct., 1930.
6. Mateer, F. Pituitary acceleration of IQ's. *Psychol. Bull.* 31: 686, Nov., 1934.
7. Mateer, F. A first study of pituitary dysfunction in cases of reading difficulty. (Abstract.) *Psychol. Bull.*, 32: 736, Nov., 1935.
8. Menninger, W. C. The mental disturbances associated with pituitary disorders. *J. A. M. A.*, 91: 951–955, Sept. 29, 1928.
9. Novak, E. Anterior pituitary and anterior pituitary-like substances: Therapeutic applications. Glandular physiology and therapy: A symposium. Chapter VIII, pp. 119–138, A. M. A., Chicago, 1935.
10. Potter, H. W. Endocrine therapy in mental deficiency. *Endocrinology*, 7: 25–40, Jan., 1923.
11. Potter, H. W. Résumé of research work at Letchworth Village. *Ment. Hyg.*, 9: 772–782, Oct., 1925.
12. Rubinstein, H. S. The production of testicular descent with the water soluble (anterior pituitary-like) fraction of pregnancy urine. *Endocrinology*, 18: 475–481, July–August, 1934.
13. Schott, E. L. Superior intelligence in patients with nervous and mental illnesses. *Jr. Abnor. and Social Psychol.*, 26: 94–101, Apr.–June, 1931.
14. Schott, E. L. School maladjustments of some mentally superior patients in a psychiatric clinic. *Psychol. Clinic*, 21: 202–207, Sept.–Nov., 1932.

AN EXPERIMENT TO TEST THE RELIABILITY OF ESTIMATES OF USE OF TIME

MAPHEUS SMITH
University of Kansas

AS an accompaniment of an increase in leisure time and the problem of its use it is to be expected that research in the field also will increase. Another reason for an increase in such studies is the obvious relationship between manner of using time and one's social and personal characteristics. If a complete record were available of the exact use of time and the specific contacts and associations of any person in the usual routine of his life, the information would provide an excellent objective basis for interpreting significant aspects of his personality and social relationships. This factor and the human interest of such studies has generally prevailed over scientific caution with the result that studies based on estimates of use of time have thus far outstripped tests of the reliability of the estimates employed in the studies.

The present investigation of the reliability of estimates of use of time cannot claim to give a complete answer to even one question about reliability of such estimates, but at least it provides some significant information on one of the simpler questions: the reliability of use of time for a period of one week. By determining the reliability of each item estimated for such a brief time a definite gain is made, because most items which cannot be reliably estimated for a short time are hardly likely to be reliable for longer periods.[1]

[1] This is not necessarily true of all items, however. Many people may have cycles of activity that run on the basis of two weeks or a month, rather than on a weekly basis. Weekly estimates, therefore, may not be "unusual," and yet may be sharply different from recorded times. The same person might make much more accurate estimates on a longer time basis than on a shorter one. Yet such people and the items concerned are probably in the minority.

The 81 subjects employed in this experiment were students taking first courses in sociology at the university level. The subjects were requested to make considered estimates of the amount of time they spent in each of twenty-four activities during an average week for the period of the year of the experiment (two groups in the last month of the spring semester, one group in the first portion of one summer session). The students were not permitted to retain copies of their estimates. During the ten days after the estimate was made each student kept a daily record of the use of his time for a full calendar week. The directions stated that a few minutes before retiring should be set aside each evening during the week to record the expenditures of time. Only the records of subjects who expressed themselves as entirely willing to coöperate were used, and who at the conclusion of the experiment reported that the week was not particularly unusual.

The following items were included in the study:

1. Sleeping—Exclusive of reading, study, or any other items in the list.
2. Eating—Time engaged in eating; exclusive of after dinner conversation.
3. In class—Includes lectures, laboratory, and recitation periods.
4. Preparation—Study on term papers, or other assigned work for college courses.
5. Reading magazines—Periodical literature not included in preparation.
6. Reading books—Printed books not included in preparation.
7. Reading newspapers—Daily or weekly newspapers not included in preparation.
8. Motion pictures—Time spent in attendance.
9. Music—Time spent in listening to or playing music and when no other use is being made of the time.
10. Religious activities—Includes attendance at religious services, but no reading or preparation for college courses, such as Bible Study.
11. Dancing—Does not include time spent listening to music.
12. Card playing—Includes group games and solitaire.
13. Other non-athletic games—Confined to indoor games of

a less vigorous nature, such as pool and billiards, backgammon, checkers, and chess, but not bowling.

14. Athletic games—Includes vigorous indoor games and outdoor sports.
15. "Dates"—Time spent with single members of the opposite sex.
16. Riding and walking for pleasure—Includes motoring, bicycling, horseback riding, and hiking.
17. Housekeeping—Includes housecleaning, putting one's room in order, cooking.
18. Personal appearance—Includes bathing, dressing, etc.
19. Transportation—Includes time spent in getting to and from classes, and other places where time is spent, including church, motion picture houses, boarding club, and residence.
20. Letter writing—Includes personal and business correspondence.
21. Extra-curricular activities—Limited to time spent with non-fraternity organizations and at special lectures and forums, etc.; does not include time under any other items of this list.
22. Conversation—Includes only time when conversation is the chief or only activity; does not include time counted under "music," "dancing," "card-playing," "dates," etc.
23. Work—Includes only time of employment for wages or salary.
24. "Doing nothing"—Complete idleness included here; exclusive of miscellaneous time or time unaccounted for.

Careful consideration will show that this list does not include every activity of college students, for example, parties and fraternity activities. Also some of the items are still rather complex. However, the items that have been included embrace the most time-consuming and definite uses of time of college students, as well as of the general public.

Two methods of determining the reliability of estimates were used: 1. percentage differences between estimated and recorded time in terms of recorded time, and 2. coefficient of correlation of estimate and record. According to the first method estimates of the amount of time spent in sleeping were

TABLE I

Comparison of Students' Estimates and Records of Use of Time for One Week

	AVERAGE NO. OF HRS.		DIFFERENCE BETWEEN ESTIMATES AND RECORDS (+ INDICATES OVER-ESTIMATE)	$\frac{\text{DIFF.}}{\text{RECORD}}$ %	COEFFICIENT OF CORRELATION OF ESTIMATES AND RECORDS
	Estimate	Record			
1. Sleeping	52.54	51.88	+0.66	1.3	.59 ± .05
2. Eating	8.74	8.40	+0.34	4.0	.71 ± .04
3. In class	16.72	15.57	+1.15	7.4	.79 ± .03
4. Preparation	28.11	27.61	+0.50	1.8	.78 ± .03
5. Reading magazines	1.53	1.42	+0.11	7.8	.49 ± .06
6. Reading books	1.69	1.18	+0.51	43.2	.59 ± .05
7. Reading newspapers	2.74	3.09	0.35	11.3	.66 ± .04
8. Motion pictures	1.97	1.65	+0.32	19.4	.53 ± .06
9. Music	1.82	1.56	+0.26	16.7	.49 ± .06
10. Religious activities	0.99	0.86	+0.13	15.1	.76 ± .03
11. Dancing	1.41	1.11	+0.30	27.0	.75 ± .03
12. Card-playing	1.35	1.12	+0.23	20.5	.59 ± .05
13. Other non-athletic games	0.26	0.21	+0.05	23.8	.75 ± .03
14. Athletic games	2.37	1.94	+0.43	22.2	.64 ± .04
15. "Dates"	3.46	3.65	0.19	5.3	.78 ± .03
16. Riding and walking for pleasure	2.27	2.89	0.62	21.5	.47 ± .06
17. Housekeeping	3.89	4.99	1.10	22.0	.83 ± .03
18. Personal appearance	6.18	6.76	0.58	8.6	.62 ± .05
19. Transportation	5.20	5.56	0.36	6.5	.51 ± .06
20. Letter-writing	1.82	1.67	+0.15	9.0	.35 ± .07
21. Extra-curricular activities	0.93	1.37	0.44	32.1	.63 ± .05
22. Conversation	7.46	8.43	0.97	11.5	.51 ± .06
23. Work	4.74	4.64	+0.10	2.2	.79 ± .03
24. "Doing nothing"	7.64	5.81	+1.83	31.5	.46 ± .06
TOTAL	165.83	163.37	2.46		
Time unaccounted for[a]	2.17	4.63	2.46		
Total hours in week	168.00	168.00			

[a] A sum of picnics and parties, fraternity activities, and inaccuracy of records or incompleteness of estimates.

the most accurate (1.3 per cent difference), followed by time spent in preparation (1.8 per cent difference). See Table I. Greatest inaccuracy was found in estimates of time spent in reading books (43.2 per cent difference), and extra-curricular activities (32.1 per cent difference). In addition to time spent in sleeping and in preparation, items for which fairly accurate estimates were made include eating, time in class, time spent reading magazines, time spent during "dates," time spent on personal appearance, time spent on transportation, time spent in writing letters, and time spent in gainful work. In addition to time spent reading books, in extra-curricular activities, and in "doing nothing," items in which there was more than a 20 per cent error were dancing, card-playing, other non-athletic games, athletic games, riding and walking for pleasure, and housekeeping. None of these activities occurs with any degree of regularity, with the exception of extra-curricular activities and housekeeping, and even the first of these is not always possible to reduce to an easily predicted schedule.

The case of housekeeping, as well as some other items, demonstrates the shortcomings of the first method of study of reliability. The percentage difference between the average estimated time and average recorded time reveals little about the correlation or co-variation of the estimates and records. In some items the correlation coefficient was sharply at variance with the percentage difference, for example, housekeeping, which had the highest coefficient of correlation and was among the items showing highest percentage difference. Other examples include sleeping which had a relatively low correlation, reading magazines, letter writing, music, religious activities, and other non-athletic games. Although some of these coefficients seem very inconsistent with the percentage difference figures, the coefficients must take precedence as a technique of determining reliability, because they do take account of variations as well as averages, are measures of simi-

larity of record and estimate for each subject and do not conceal individual errors of estimate in ordinary group averages.

According to the strictest standards of interpretation of the correlation coefficient—that only coefficients above .90 have real significance[2]—none of the items gave reliable results. A coefficient of this magnitude is only a little better than half way between a perfect correlation and a guess. If a somewhat less strict standard of interpretation is employed—for example, that a coefficient of correlation is high if above .70—, estimates of time spent in eating, in class, in preparation, in religious activities, in dancing, in other non-athletic games, in "dates," in housekeeping, and at work were reliable enough for use in future studies. Reliability of the estimates on several of these items is not surprising, because of the definite habits they represent. This is true of time spent in eating, in class, in preparation, in religious activities, and at work. On the other hand, the reliability of estimates of time spent in dancing, in other non-athletic games, and in "dates" may have been due to chance. Certainly there is nothing in the nature of these items that would make them easier to estimate than sleeping, transportation, and personal appearance.

Analysis of the records on the last three items show that relatively few subjects who made especially poor estimates accounted for the low coefficients. In sleeping, for instance, one subject estimated 64 hours for the week and slept only 50 hours, and for a number of others the estimate failed by more than ten hours to coincide with the record. For transportation individual errors of estimate of as much as ten hours were found, although the averages of both estimated and recorded time were less than six hours for the entire group of subjects. And in the case of personal appearance differences were as large as 7.5 hours in some cases. Some of these extreme variations may be due to the various activities, and some may be due to erroneous statements about the degree of usualness of

[2] Garrett, H. E., *Statistics in Psychology and Education*, New York, 1926, p. 290.

the use of time during the experimental week.[3] Another factor may have been the variability of use of time over a period greater than one week, which would make it quite impossible to estimate some activities accurately on the basis of one week, and which would also make incorrect the claim that the week was unusual or that the subject had been careless or was not intelligent.

Items of most doubtful reliability were time spent in letter-writing, in walking and riding for pleasure, in playing or listening to music, in reading magazines, and in "doing nothing." Because of the unstereotyped nature of these uses of time it is not surprising that estimates were often wide of the mark. The same thing is true of time spent in conversation, which had one of the lower coefficients.

There are numerous implications of this study for further research in the reliability of estimates of use of time. In the first place, the study is too small to justify broad generalizations even for estimates of use of time for one week. More extensive studies on the same and different ranges of time and at different periods of the year should be conducted. A more careful categorizing of use of time including more clear-cut definitions and more exact directions should be employed. Other types of subjects should eventually be studied. This is particularly important because, although college life is highly organized at certain definite points, use of time probably is more stereotyped in both earlier and later years. Such methodological experiments should be tried out on groups of

[3] Where the subject considered the week to be especially unusual, the record was not included in the final tabulations. Discard of subjects because of the unusual expenditures of time is obviously a crucial point in reliability studies. To have any fair measure of reliability obviously worthless cases must be eliminated. But if all variant cases are omitted, the reliability is probably increased unjustifiably. In this study the subject was asked if the week were especially unusual in some item or items that might affect the general estimate. The fact that only a few subjects expressed the opinion that the week of the experiment was unusual may be one reason the estimates were of such low reliability.

people with as widely different life organizations as possible. And of paramount importance, the subjects should be stimulated to coöperate to the fullest in the experiment, both in estimating and in keeping records. The procedure, to insure greatest success, would involve careful supervision and even interviews to check the reasoning back of the estimate, upon the accuracy of which this method of research ultimately depends. It is possible that the extreme unreliability of many procedures and techniques is due for the most part to the lack of care in stimulating and aiding the subjects to do their best. Although many types of information are inherently unreliable because of their subjectivity, the infrequency of the subjects' experience with the facts desired, and the lack of intelligence of subjects, it may be that persons with sufficient ability to govern their ordinary affairs can make somewhat more reliable estimates of use of time than those reported when all of these safeguards are employed. If the results presented above are proved by further investigations to be a true indication of the reliability of estimates of use of time, the method of estimate is highly unsatisfactory for serious research. However, the tediousness of keeping detailed records which may be done for short experimental periods, or for longer periods for fewer subjects probably will prevent the use of records of use of time for extensive studies. It is, therefore, of considerable importance to determine if any means can be discovered for improving the reliability of estimates for use in future studies.

PSYCHO–PHYSIOLOGICAL REACTIONS FOLLOWING PAINFUL STIMULI UNDER HYPNOTIC ANALGESIA CONTRASTED WITH GAS ANAESTHESIA AND NOVOCAIN BLOCK

RALPH R. BROWN

Psychologist, United States Public Health Service

AND

VICTOR H. VOGEL

Past Assistant Surgeon, United States Public Health Service

THE experiments described in this paper represent an attempt to secure a better understanding of the phenomenon of hypnotically induced analgesia through a controlled study of the involuntary and semivoluntary psychophysiological processes which are usually altered as the result of painful stimuli. We were interested in learning to what extent physiological reactions to painful stimuli under hypnotic analgesia might simulate the type of reactions occurring after a local block or general anaesthetic. We wished to know if suggested analgesia in the trance state differed significantly and consistently from waking voluntary attempts at self-induced analgesia—that is, to what extent the effects of suggestion under hypnosis may transcend, or vary from, the normal waking effect of volition upon those processes generally considered to be involuntary psycho-physiological functions. Hull (2) has called attention to the need of a study in which hypnotically induced analgesia is compared with a chemically induced anaesthesia.

Continuous tracings of blood-pressure (relative), pulse rate, skin-potential and respiratory changes, as well as movements of the right and left hands were secured by means of the Darrow Photopolygraph (1). A spring, automatic blood lancet, set at

4 mm., was used for producing sharp pain. Continuous pain was induced by means of a thumb-tack held down by a leather strap weighted at both ends. The tack was sewed into the center of the strap and was placed on the bridge of the foot, the weighted ends of the straps hanging down on each side. Total weight of the strap was 550 grams.*

Three subjects were selected as being most suitable for the present experiment. All showed post-hypnotic amnesia and from all external signs, the subjects appeared to develop a deep anaesthesia upon suggestion in the trance state. Our interest was not in observing how the 'typical' or 'average' hypnotic subject reacted to suggestion of anaesthesia, but rather we were interested in learning the differences between a well established hypnotic analgesia and other types of analgesia. Therefore our cases were selected for depth of apparent analgesic effects under hypnosis, and consequently it must be borne in mind that we are not dealing here with the average hypnotized subject.

Four preliminary experiments were made on each subject, two in the waking state and two in the trance state, after which the main series of experiments were started. These preliminary experiments were intended not only to accustom the subject to the apparatus and rid him of anticipatory excitation, but also served to furnish us with a knowledge of the general reactance level of each subject and to give us an idea as to the efficacy of our main experimental program. These preliminary experiments were concerned with physiological reactions to sensory and ideational stimuli in the trance and waking states. *It was quite apparent from these experiments that suggested deafness in the trance state does not abolish physiological reactions to auditory stimuli. It was further established that suggested anaesthesia does not abolish reactions to tactile and pain stimuli.*

The main experimental program consisted of observations on the physiological reactions to sharp pain (lancet), continuous pain (thumb-tack), and hot water (49 degrees centigrade) under the following conditions:

* Strap was designed by Victor H. Vogel.

1. During normal waking state
2. Under hypnosis with suggestion of complete analgesia
3. During normal waking state with request to imagine analgesia
4. Under general gas anaesthesia
5. After novocain block
6. During hypnosis with suggestion of hyperalgesia
   a. Effects of painful stimuli
   b. Effects of indifferent stimuli
7. During normal waking state with request to imagine hyperalgesia
   a. Effects of painful stimuli
   b. Effects of indifferent stimuli

The pain stimuli were administered to the right foot. The left foot was used as a control during the experiments on novocain block, hyperalgesia, and analgesia. The order of the experiments was different for each subject. In the experiments on hyperalgesia, both painful and indifferent stimuli were used. As indifferent stimuli, a blunt pencil point was used in place of the lancet, the flat button of the thumb-tack was used for suggestion of continuous pain, and tepid water was used in place of the hot water. During the experimental runs under hypnosis, Dr. Vogel gave the instructions (including induction of hypnosis), whereas during the waking runs, Dr. Vogel was not present.

Results: In table 1, the general method of analyzing the results is shown. The galvanic skin response has been analyzed with respect to conductance change (expressed in micromhos and the time for recovery. The recovery ratio was calculated by dividing the conductance recovery in five seconds by the conductance change following the stimulus, times one hundred.

The effect of stimuli upon the pulse rate was figured by dividing the number of beats in twenty seconds after the stimulus by the number in twenty seconds before the stimulus, times one hundred. Changes in relative blood pressure are expressed in terms of the scale units on the record (2 mm. steps). The

TABLE # 1

EXPERIMENT	STIMULUS	SUBJECT	CONDUCTANCE CHANGE			RECOVERY RATIO			PULSE RATE AFTER − BEFORE × 100			BLOOD PRESSURE UNITS CHANGE			TIME for 5 BREATHS (sec)			EXP. AMP / INSP. AMP × 100		
NORMAL WAKING STATE	HOT WATER	G	000	02	+	01	000	+	100	104	+	50	1.25	+	14.50	14.00	−	100.0	100.0	0
		J	00	02	+	00	00	0	100	100	0	.75	1.00	+	14.25	13.00	−	116.7	116.7	0
		Mc	26	29	+	19	17	−	104	103	−	1.50	.25	−	30.50	23.75	−	100.0	100.0	0
	LANCET	G	12	09	−	08	11	+	104	104	0	1.75	.50	−	15.25	15.75	0	92.5	91.7	+
		J	35	28	+	18	09	−	96	96	0	1.25	.25	−	14.50	13.50	−	116.7	100.0	−
		Mc	24	27	+	79	28	−	100	100	0	3.00	.75	−	18.75	20.00	+	118.5	91.7	−
	THUMB TACK	G	24	30	+	31	11	−	98	107	+	1.75	2.00	+	12.50	12.75	+	100.0	100.0	0
		J	29	37	+	07	27	+	100	98	−	.50	1.25	+	14.75	14.75	0	100.0	112.5	+
		Mc	20	28	+	25	37	+	109	100	−	2.25	2.00	+	19.00	20.00	+	100.0	100.0	0
WAKING ANALGESIA	HOT WATER	G	00	00	0	00	00	0	104	104	0	.00	.75	+	12.50	12.00	−	100.0	109.1	+
		J	03	08	+	33	18	−	104	92	−	1.00	1.25	+	11.75	12.75	+	100.0	100.0	0
		Mc	04	11	−	00	00	0	105	105	0	2.00	.50	−	21.00	20.00	−	123.1	111.1	+
	LANCET	G	27	15	−	11	08	−	120	93	−	1.00	2.00	+	14.50	13.00	−	100.0	88.9	−
		J	71	47	−	18	14	−	100	92	−	3.00	3.25	+	11.75	13.00	+	100.0	100.0	0
		Mc	58	33	+	78	19	−	100	91	−	1.00	.50	−	21.50	24.50	+	111.1	100.0	−
	THUMB TACK	G	31	30	−	16	07	−	114	115	+	1.50	3.00	+	††	14.00		††	92.5	
		J	69	54	−	10	20	+	98	96	−	.25	1.00	+	12.75	13.50	+	100.0	100.0	0
		Mc	49	32	−	22	69	+	100	95	−	.75	.50	−	18.00	26.50	+	100.0	100.0	0
WAKING HYPERALGESIA	TEPID WATER	J	00	05	+	00	40	+	98	99	+	1.00	2.25	+	14.25	14.00	−	100.0	100.0	0
		Mc	11	48	+	00	10	+	100	100	0	.50	.25	−	22.25	22.75	+	100.0	96.2	−
INDIFFERENT STIMULI	DULL POINT	G	03	04	+	00	50	+	94	100	+	1.00	2.00	+	14.75	13.25	−	100.0	100.0	0
		J	↳→																	
		Mc	07	11	+	28	00	−	96	100	+	.25	.50	+	24.75	23.75	−	100.0	100.0	0
	THUMB TACK UPSIDE DOWN	G	02	01	−	50	00	−	97	98	+	1.00	1.25	+	14.50	15.00	+	100.0	100.0	0
		J	↳→																	
		Mc	10	38	+	00	00	+	100	100	0	1.50	1.25	−	22.00	21.00	−	100.0	100.0	0
WAKING HYPERALGESIA	HOT WATER	G	02	17	+	00	00	0	98	103	+	1.00	.75	−	10.50	12.00	+	105.6	103.3	−
		J	↳→																	
		Mc	00	24	+	00	**		100	100	0	.25	1.50	+	23.50	24.00	+	107.1	100.0	−
	LANCET	G	↳→	04	−	00	00	0	103	97	−	50	2.75	+	15.00	13.00	−	100.0	100.0	0
PAINFUL STIMULI		Mc	47	49	+	03	08	−	108	92	−	2.50	5.00	+	23.00	24.50	+	92.3	100.0	+
	THUMB TACK	G	22	32	+	09	04	−	103	104	+	.50	3.50	+	15.00	17.00	+	95.5	100.0	+
		J	↳→																	
		Mc	48	34	−	13	04	−	100	96	−	1.75	1.75	0	19.75	20.50	+	104.2	104.5	−
HYPNOTIC ANALGESIA	HOT WATER	G	01	01	0	00	00	0	103	100	−	1.00	1.00	0	12.75	13.50	+	100.0	100.0	0
		J	11	44	+	09	23	+	96	102	+	.75	1.00	+	11.75	13.50	+	125.0	100.0	−
		Mc	10	15	+	00	33	+	95	90	−	2.25	1.50	−	14.00	20.25	+	110.0	125.0	+
	LANCET	G	13	08	−	11	**		100	98	−	1.75	1.50	−	21.00	12.50	−	108.7	92.3	−
		J	138	122	−	23	27	+	104	96	−	1.50	.25	−	14.25	13.50	−	100.0	125.0	+
		Mc	32	13	+	61	24	−	104	98	−	2.50	1.00	−	21.75	23.75	−	100.0	108.7	+
	THUMB TACK	G	27	12	−	19	29	+	††	103		2.25	3.00	+	10.25	14.50	+	100.0	100.0	0
		J	56	16.5	+	19	29	+	98	94	−	2.00	1.00	−	14.00	14.50	+	100.0	100.0	0
		Mc	30	24	−	25	120	+	96	83	−	2.00	1.00	−	17.50	35.00	+	100.0	100.0	0
HYPNOTIC HYPERALGESIA	TEPID WATER	G	05	04	+	00	25	+	98	97	−	1.50	2.25	+	13.50	13.00	−	100.0	100.0	0
		J	00	07	+	00	29	+	100	97	−	.50	0.00	−	13.00	12.50	−	100.0	100.0	0
		Mc	28	28	0	50	100	+	100	98	−	1.00	3.00	+	28.25	29.50	+	100.0	100.0	0
INDIFFERENT STIMULI	DULL POINT	G	00	01	+	00	00	0	102	97	−	3.00	3.00	0	14.00	14.00	0	100.0	100.0	0
		J	14	08	−	00	00	0	100	100	0	1.00	1.50	+	13.25	13.25	0	100.0	100.0	0
		Mc	114	43	−	28	21	+	104	96	−	2.00	4.00	+	20.00	18.50	−	100.0	100.0	0
	THUMB TACK	G	2**	10	+	00	**		102	113	+	.50	1.00	+	14.00	14.00	0	100.0	100.0	0
	UPSIDE DOWN	J	20	39	+	05	10	+	98	102	+	1.50	1.50	0	13.25	13.50	+	100.0	100.0	0
		Mc	37	30	−	70	63	−	97	96	−	1.00	**		41.50	52.50	+	98.1	100.0	+
HYPNOTIC HYPERALGESIA	HOT WATER	G	00	04	+	00	25	+	100	103	+	2.00	2.00	+	13.00	16.25	+	100.0	100.0	0
		J	00	00	0	00	00	0	100	98	−	2.00	2.75	+	15.00	13.50	−	100.0	100.0	0
		Mc	50	48	−	10	13	+	107	100	−	8.00	39		45.25	60.75	+	100.0	100.0	0
PAINFUL STIMULI	LANCET	G	9	09	−	00	00	0	101	102	+	1.00	1.00	0	13.00	13.00	0	100.0	81.8	−
		J	**	3***		**	**		97	102	+	2.00	3.00	+	14.00	13.75	−	100.0	100.0	0
		Mc	73	75	0	39	39	0	100	96	−	2.25	1.50	−	49.50	41.00	+	100.0	100.0	0
	THUMB TACK	G	**21	22	−	**	00		97	100	+	1.00	**		10.50	23.25	+	100.0	93.5	−
		J	7	37	+	14	16	+	97	98	+	.50	.25	−	15.00	13.75	−	100.0	100.0	0
		Mc	33	52	+	50	98	+	98	98	0	††	1.50		30.25	39.00	+	100.0	103.1	+
LOCAL BLOCK	BLUNT POINT	G	9	06	−	33	67	+	99	97	−	1.25	1.00	−	16.50	17.00	+	100.0	100.0	0
		J	4	00	−	00	00	0	98	102	+	.50	.75	+	16.00	15.00	−	100.0	114.3	+
		Mc	47	35	−	70	31	−	96	88	0	3.00	1.00	−	30.75	27.00	−	100.0	94.1	−
INDIFFERENT STIMULI	THUMB TACK	G	↳→																	
		J	14	14	+	33	21	−	100	100	0	2.00	1.00	−	19.50	16.75	−	88.2	100.0	+
	UPSIDE DOWN	Mc	↳→																	
LOCAL BLOCK	LANCET	G	9	08	−	33	38	+	97	98	−	2.00	1.00	−	16.75	16.25	−	100.0	100.0	0
		J	52	25	−	15	08	−	94	96	+	3.00	.50	−	15.25	15.75	+	100.0	114.3	+
		J	11	11	0	0	18	+	102	98	−	2.25	.75	−	14.75	13.50	−	100.0	114.3	+
		Mc	35	63	+	88	28	−	104	100	−	1.50	3.00	+	22.25	27.00	+	100.0	100.0	0
PAINFUL STIMULI		Mc	74	84	+	43	93	+	105	96	−	1.75	2.50	+	45.25	42.25	−	100.0	100.0	0
	THUMB TACK	G	14	09	−	33	33	0	100	98	−	2.00	1.00	+	14.75	13.75	−	100.0	108.7	+
		J	11	11	+	00	09	+	96	96	0	.50	1.00	+	15.75	16.00	+	100.0	100.0	0
		Mc	37	33	+	00	25	+	105	100	−	2.75	.50	−	20.00	20.50	+	100.0	94.1	−
GAS ANAESTHESIA	HOT WATER	G	§§	0		§§	0		§§	89		§§	50		§§	14.00		§§	106.0	
		J	***↳																	
		Mc	0	§§		0	§§		100	§§		2.50	§§		20.00	§§		100.0	§§	
	LANCET	G	0	0	0	0	0	0	99	100	+	50	50	0	14.25	15.50	+	100.0	107.0	+
		J	***↳																	
		Mc	§§	0		§§	0		§§	95		§§	3.00		§§	11.50		§§	100.0	
NITROUS OXIDE	THUMB TACK	G	§§	0		§§	0		§§	104		§§	50		§§	27.75		§§	103.0	
		J	***↳																	
		Mc	§§	0		§§	0		§§	103		§§	1.00		§§	15.50		§§	100.0	

* LEFT, CONTROL FOOT † RIGHT, EXPERIMENTAL FOOT. ‡ EFFECT PLUS INDICATES GREATER PHYSIOLOGICAL EFFECT ON EXPERIMENTAL (RIGHT) FOOT. § SEVERE FATIGUE STATE NOTED AFTER COMPLETION OF RECORD. DATA NOT USED

** GALVANOMETER WENT OFF RECORD †† SUBJECT UNABLE TO KEEP THUMB TACK ON FOOT. ‡‡ SUBJECT MOVED ARM CAUSING DISTURBANCE IN BLOOD PRESSURE RECORDING. §§ STIMULUS NOT GIVEN. *** SUBJECT WAS NOT ANAESTHETIZED

respiration was analyzed with respect to time for five breaths, as well as with respect to the relationship between expiration and inspiration amplitude. The I/E ratio was not calculated.

Among other things, Table 1 shows:

1. Marked variability of reaction to painful stimuli both with respect to extent and to direction of effect. Although the stimuli were judged to be quite painful, there seemed to be no physiological signs of sufficient significance for determining the presence of pain.

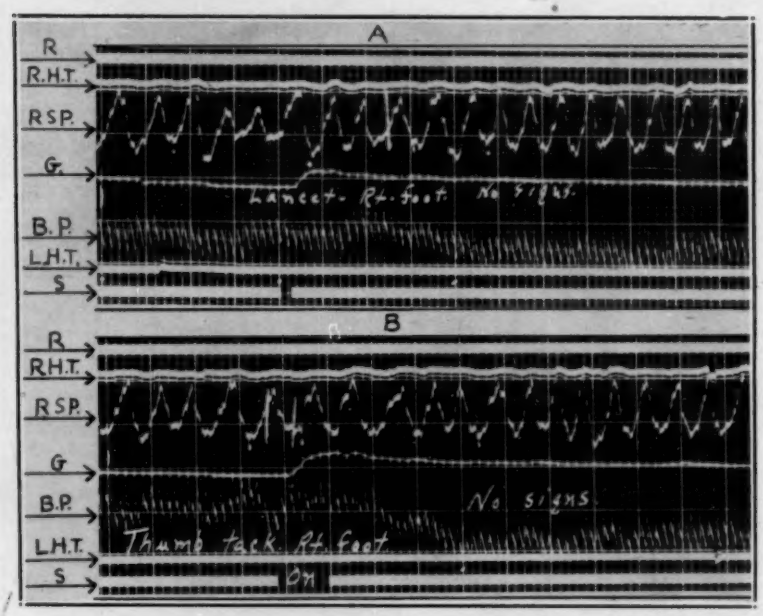

FIGURE 1. A: Polygraphic picture of changes in *normal waking state* following application of sharp pain (Lancet) to right foot. Subject Mc. Figure 1. B: Polygraphic picture of changes in *normal waking state* following application of continued pain (thumb tack) to right foot. Subject Mc.

Legend: R. Response marker; R.H.T. Right hand tremor; R.S.P. Respiration. Up line, Inspiration. Down line, Expiration; G. Galvanic skin reflex; B.P. Blood pressure from left arm; L.H.T. Left hand tremor; S. Stimulus marker.

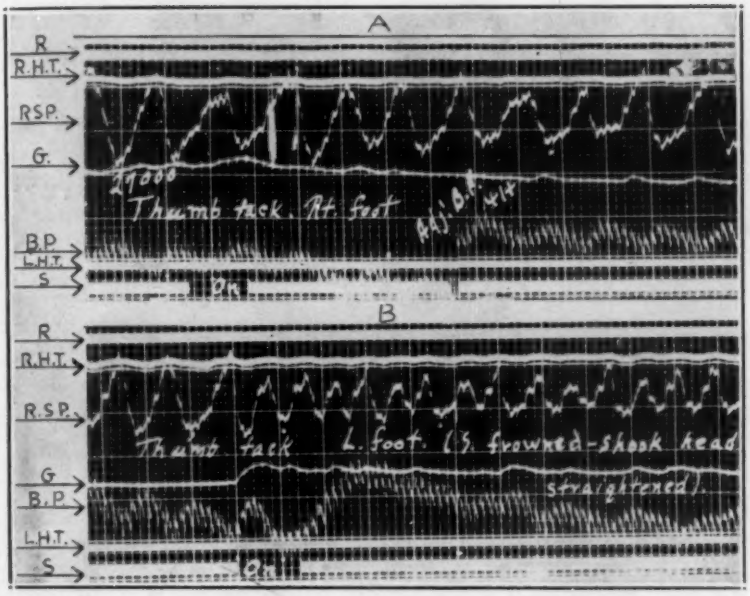

FIGURE 2. A: Polygraphic picture of changes in *trance state* following application of continued pain (thumb tack) to right foot with suggestion of analgesia. Subject Mc. Figure 2. B: Polygraphic picture of changes in *trance state* following application of continued pain (thumb tack) to left foot with suggestion of normal feeling. Subject Mc.

Legend: R. Response marker; R.H.T. Right hand tremor; R.S.P. Respiration. Up line, Inspiration. Down line, Expiration; G. Galvanic skin reflex; B.P. Blood pressure from left arm. L.H.T. Left hand tremor; S. Stimulus marker.

2. There appears to be no marked difference between the physiological reactions to pain under hypnotically induced analgesia and that under imagined analgesia in the normal waking state, although only a tentative conclusion can be made on the basis of our present study. From general observation of the patients, however, it appeared that they were much more willing or able to withstand pain (no external signs) in the hypnotic state. One subject in the waking state repeatedly kicked the thumb-tack off his foot, but allowed the tack to remain on after being requested to do so.

3. Gas anaesthesia is not comparable to hypnotic anaesthesia. The former induces a marked change in the physiological state which is quite dissimilar to that observed during hypnotic anaesthesia. In one subject (Q), the blood pressure rose from 106/74 to 195/100, with the pulse rate changing from 79 to 117.

Figure 1 shows the physiological reactions of subject Mc to lancet and thumb-tack during the normal waking state. It may be noted that there is no indication of a severe disturbance of the physiological functions. There is a slight rise in blood

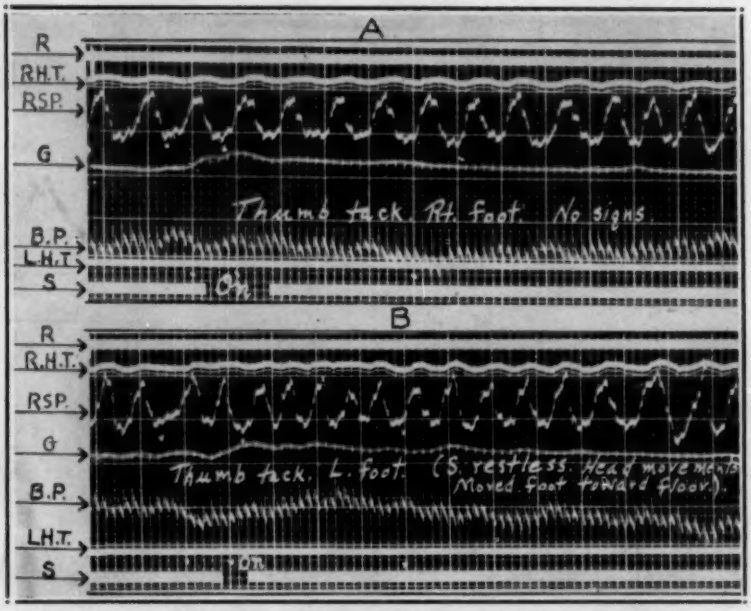

FIGURE 3. A: Polygraphic picture of changes in *waking state* following application of continued pain (thumb tack) to right foot with suggestion of analgesia. Subject Mc. Figure 3. B: Polygraphic picture of changes in *waking state* following application of continued pain (thumb tack) to left foot with suggestion of normal feeling. Subject Mc.

Legend: R. Response marker; R.H.T. Right hand tremor; R.S.P Respiration. Up line, Inspiration; Down line, Expiration; G. Galvanic skin reflex; B.P. Blood pressure from left arm; L.H.T. Left hand tremor; S. Stimulus marker.

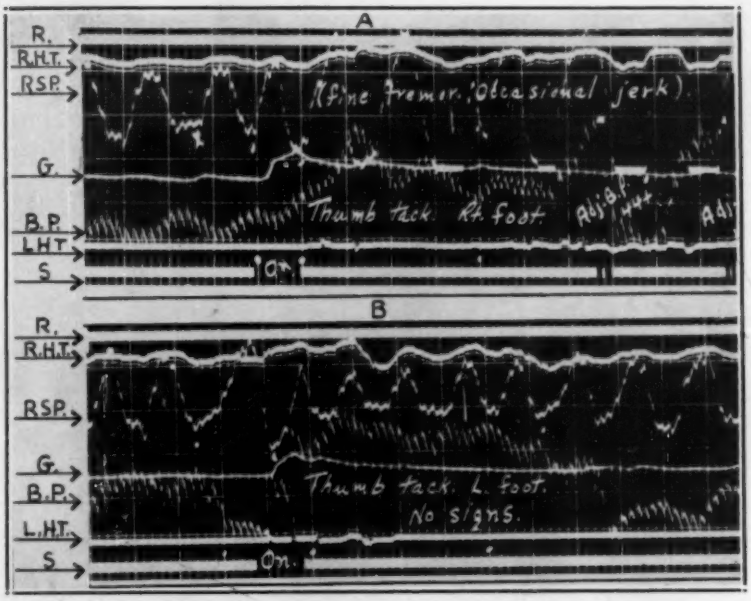

FIGURE 4. A: Polygraphic picture of changes in *trance state* following application of continued pain (thumb tack) to right foot with suggestion of hyperalgesia. Subject Mc. Figure 4. B: Polygraphic picture of changes in *trance state* following application of continued pain (thumb tack) to left foot with suggestion of normal feeling. Subject Mc.

Legend: R. Response marker; R.S.P. Respiration. Up line, Inspiration. Down line, Expiration; G. Galvanic skin reflex; B.P. Blood pressure from left arm; L.H.T. Left hand tremor; S. Stimulus marker.

pressure, a slight increase in tremor, and there is a moderate skin reflex. There is a slight increase in pulse rate following application of the thumb-tack, but no increase from the lancet. After the subject has been hypnotized and analgesia has been suggested, as shown in figure 2-A, the amplitude of the skin reflex is not as great, but there is a marked increase in the number of spontaneous fluctuations. The blood pressure tends to drop gradually, and the tremor line is quite steady. When the thumb-tack is applied to the left foot, with the suggestion of

normal feeling, however, as may be noted in figure 2-B, there are marked changes in respiration, blood pressure, and skin resistance, as well as external signs of pain. These physiological changes are in line with what would be expected from definite painful stimuli, and the extent of effect is much greater than that which occurred in the normal waking state.

In the normal waking state, when the subject is asked to imagine that his right foot is analgesic, the record is quite steady (figure 3-A). In the control record, however, there is a definite change in the respiration, blood pressure, and skin

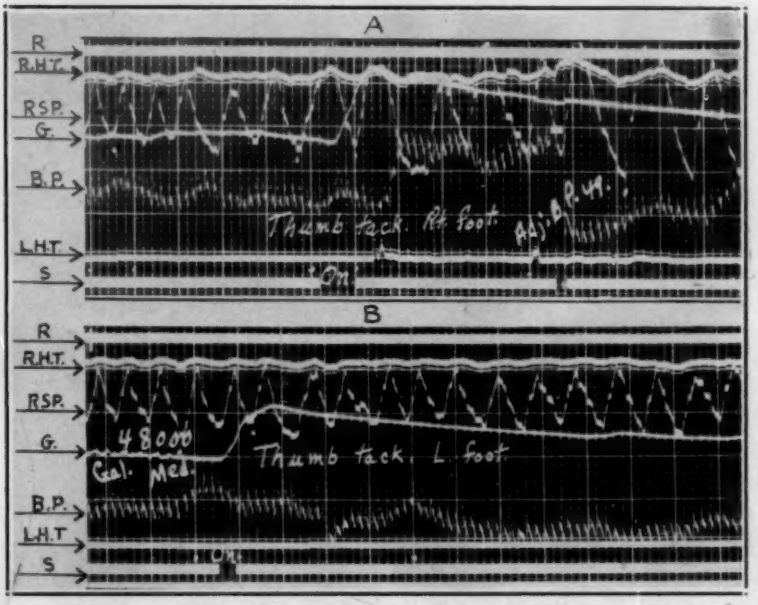

FIGURE 5. A: Polygraphic picture of changes in *waking state* following application of continued pain (thumb tack) to right foot with suggestion of hyperalgesia. Subject Mc. Figure 5. B: Polygraphic picture of changes in *waking state* following application to continued pain (thumb tack) to left foot. Subject Mc.

Legend: R. Response marker; R.S.P. Respiration. Up line, Inspiration. Down line, Expiration; G. Galvanic skin reflex; B.P. Blood pressure from left arm; L.H.T. Left hand tremor. S. Stimulus marker.

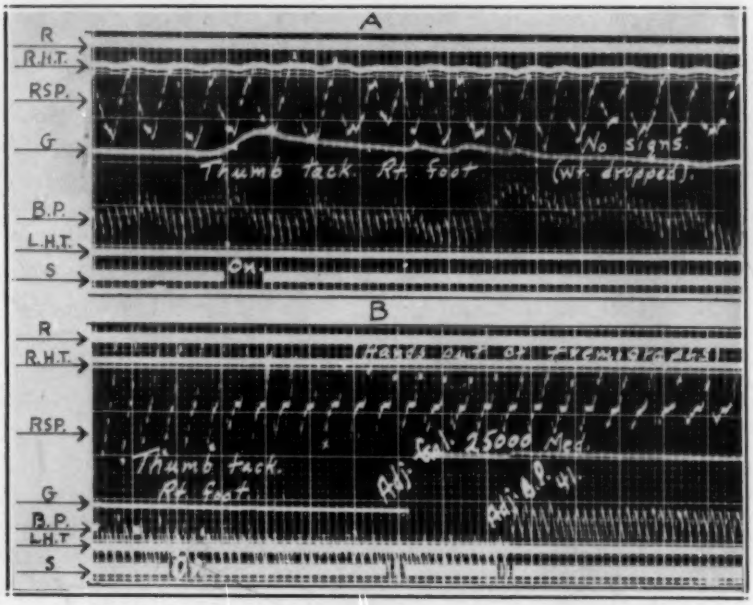

FIGURE 6. A: Polygraphic picture of changes in *waking state* following application of continued pain (thumb tack) to right foot after local novocain block. Subject Mc. Figure 6. B: Polygraphic picture of changes during *nitrous oxide anaesthesia* following application of continued pain (thumb tack) to right foot. Subject Mc.

Legend: R. Response marker; R.H.T. Right hand tremor; R.S.P. Respiration. Up line, Inspiration. Down line, Expiration; G. Galvanic skin reflex; B.P. Blood pressure from left arm; L.H.T. Left hand tremor; S. Stimulus marker.

resistance, together with external indications of pain. It is obvious that the physiological reactions are affected, although the changes are not as great as those observed in the trance state.

When hypersensitivity is suggested during the trance state, the physiological reactions may be markedly increased, as may be seen in the next figure (figure 4). There is a sharp rise in blood pressure, movements of the right and left hands are increased, and the respiration becomes deeper and more irregular.

In the lower record, it may be seen that the tendency to increased response is carried over to the left (control) foot, even though 'normal feeling' has been suggested.

It appears, however, that in the subjects which we have observed, imagination in the normal waking state may be just about as effective with respect to influencing the physiological reactions to pain as that which may be observed following suggestion in the trance state. In the next figure (figure 5-A), we have a record of the physiological responses to the thumbtack in the normal waking state. The subject has been requested to imagine that his right foot is especially sensitive. It may be noted that the responses are of about the same magnitude as those observed during the trance state. In this case, however, the hyperalgesia is not carried over to the control foot, as may be seen in the lower record (figure 5-B).

The effects produced by imagination in the normal waking state are of significance in demonstrating the possibility of *voluntarily* influencing those physiological processes which have generally been considered as autonomous. The finding lends support to the idea that psychological processes may play an important role in facilitating or inhibiting the physiological functions concerned with maintaining or reestablishing physical health. The conscious and voluntary influences upon physiological function should be emphasized as having clinical significance, inasmuch as the influence of unconscious activity upon these processes has perhaps been frequently overemphasized.

The influence of imagination upon the physiological processes in the subjects coming under our observation may also be of significance in indicating the nature of hypnosis, or, at least, in emphasizing an aspect of hypnosis which has been neglected in modern interpretations. We tend to look upon hypnosis at the present time as a form of suggestibility (Bernheim), perhaps a shift in prestige suggestibility (Hull), rather than as a form of sleep (Pavlow), or as dissociation (Janet). But the emphasis upon suggestibility neglects not only the voluntary

aspect of hypnosis, but also tends to overlook the apparent linkage between hypnosis and imagination.

The difficulty with considering hypnosis as a form of suggestibility is that the concept does not adequately describe the conditions essential for the induction of the trance. Our time limit prevents a full discussion of this aspect of the problem, however.

Figure 6 shows the influence of a painful stimulus upon the physiological reactions following chemically induced anaesthesia—the upper record (6-A) shows the reactions following novocain block and the lower record (6-B) shows the effects of general gas anaesthesia (nitrous oxide). It will be noticed that there is a definite skin response following application of the thumb-tack during novocain block. We were unable to eliminate the kinaesthetic cues associated with application of the thumb-tack and lancet, and consequently most of our records show a galvanic skin response to the pain stimulus. It should be mentioned that the extent of the skin response was not correlated with intensity of pain sensation. Some of the largest galvanic responses appeared while testing the area for analgesia directly following the pain stimuli. Very light application of a pin point on the blocked area caused no response (if the block was complete), but light applications of the point outside the area, or sufficiently peripheral to be appreciated as touch, caused large galvanometer deflections. With reference to respiration, blood pressure, heart rate, and tremor, however, local novocain block seems to have decreased the response to pain.

We were unable to observe any physiological responses to painful stimuli while the subjects were under gas anaesthesia. The application of nitrous oxide is associated with marked physiological changes, the nature of which is quite dissimilar to those observed following hypnotic analgesia. There appear to be no points of similarity between these two states.

Within the limits of our experimental conditions, the following conclusions appear to be justified:

1. Quantitative changes in physiological reactions to painful stimuli are not reliable indicators of painful experience.

2. Nitrous oxide anaesthesia abolishes physiological reactions to moderately intense sensory stimuli. (Severe stimuli were not used.)

3. Suggested analgesia in the hypnotic state does not abolish physiological reactions to sensory stimuli. Suggested deafness does not abolish physiological reactions to auditory stimuli.

4. Awareness of stimulation appears to be one of the influencing factors concerned in physiological response to moderate or mild stimuli.

5. Physiological reactions to moderate and mild sensory stimuli may be affected by suggestion in the hypnotic state and by imagination in the normal state.

6. *External* reactions generally associated with pain may be abolished by suggestion of analgesia under hypnosis, by request to imagine analgesia in the normal waking state, by local block, and by gas anaesthesia.

REFERENCES

1. DARROW, C. W. *The Behavior Research Photopolygraph.* J. Gen. Psy., V. 7, No. 1, July, 1932, 215–219.

2. HULL, CLARK L. *Hypnosis and Suggestibility.* 1933, page 253.

THE PART–WHOLE PROPORTION ILLUSION IN PRINTING

DONALD G. PATERSON AND MILES A. TINKER
University of Minnesota

IN connection with an inquiry regarding the effect of margins on legibility we were surprised to note that manuals of style specify that 50 per cent of a book page should be devoted to marginal space. In other words, printing specifications commit the printing industry to only a 50 per cent utilization of paper for printing purposes.

In order to discover whether or not printers follow these specifications we made a survey of printing practice. The ratio of printed area to the full page was computed for each of 200 history text books and 200 literature text books. It was found that the median book devotes 52.8 per cent of the full page to print and 47.2 per cent to margins. Although there is diversity in practice (ratios varied between .29 and .76), printers, on the average, come pretty close to following the 50 per cent white space rule.

In discussing this rule with numerous individuals, we discovered that none of them was aware of the rule. And when asked to guess the typical ratio in the average book, the estimates ran well over 50 per cent. In other words, these persons believed that a large proportion of a book page is devoted to printed matter and only a small proportion to margins. Thus it would appear that an illusion exists.

To determine the extent of this illusion we proceeded to question 928 college students. Each of these students submitted in writing his estimate in response to the following: "With reference to the typical text book or book of fiction I estimate that the percentage of the total page space devoted

to the main body of printed material (the rest of the page is composed of margins) is ——." These estimates were tabulated and the results are presented in Table I. It is obvious

TABLE I

Distribution of Estimates of 928 College Students with Reference to the Percentage of the Average Book Page Devoted to Printed Material

PERCENTAGE OF AVERAGE BOOK PAGE DEVOTED TO PRINTED MATERIAL	NUMBER OF ESTIMATES	PERCENTAGE OF ESTIMATES
90–100	53	5.7
80– 89	240	25.9
70– 79	292	31.5
60– 69	238	25.6
50– 59	71	7.7
40– 49	16	1.7
30– 39	8	0.9
20– 29	6	0.6
10– 19	4	0.4
0– 9	0	0.0
Total	928	100.0

that the overwhelming majority of these students believe that 60 per cent or more of the total page is devoted to printed matter. The most frequent estimate is in the neighborhood of 75 per cent. These results indicate an illusion effect of about 25 per cent since actual printing practice follows the 50 per cent rule.

These results suggested the desirability of attempting to measure this illusion effect experimentally.* In this study the aim was to determine quantitatively the amount of error involved in the estimation of part-whole areas uncomplicated by printed text. White cards with black centers and black cards with white centers were prepared as stimulus materials. For each kind three sizes of cards were employed, namely $4'' \times 7''$, $5'' \times 8''$ and $6'' \times 9''$. These sizes correspond with the

* We are indebted to Miss Dorothy Plain for collecting the experimental data and computing the results as a project under the writers' direction.

most frequent sizes of book pages as found in our survey of printing practice. Within each of these size groups the percentage of total card space devoted to the white or black central area varied according to five significant steps as found in the printing practice survey. The exact proportions are shown in the table of results. A total of 300 college students served as subjects. These were subdivided into three subgroups of 100 each. Each subgroup submitted estimates of the proportion of the total card size occupied by the centers for each card of a given size group. To avoid fatigue, loss of interest and a possible tendency to develop methods for mentally analyzing the problem, no subject was asked to estimate proportions in more than one size group. Each subject estimated proportions for the five cards in the white on black series and then estimated the five cards in the black on white series. The order of presentation (white on black and black on white) was varied from subject to subject in order to eliminate the effect of practice. The order of presentation of the five cards in each series was determined by chance. One card at a time in a given series was shown a subject. The subject was instructed as follows: "This experiment is concerned with the estimation of relative areas. You will be shown a series of 10 cards, one by one, and you are requested to estimate what per cent of the area of the whole card is taken up by the area in the center of the card." Three seconds was allowed for each estimation since this was found to be long enough to permit the subject to make a judgment and yet not long enough to permit the average subject to work out methods of analyzing the areas and calculating the proportions. Sex differences in estimations made by the subjects were negligible hence the results are reported for each group of subjects as a whole.

RESULTS

The results are presented in Table II. It is clear that there is a constant tendency to overestimate the proportion of the

TABLE II

Mean Estimated Proportions of Center Area to Total Card Area in Relation to Actual Proportions by Measurement

A. Size of Cards = 4″ × 7″. N = 100 Subjects

ACTUAL PROPORTIONS	BLACK ON WHITE		WHITE ON BLACK	
	Estimated Proportions	Over-estimation	Estimated Proportions	Over-estimation
42.7	59.4	16.7	64.4	21.7
49.5	66.5	17.0	67.4	17.9
53.1	70.5	17.4	72.7	19.6
59.2	78.2	19.0	79.1	19.9
71.9	86.2	14.3	88.6	16.7
B. Size of Cards = 5″ × 8″. N = 100 Subjects				
38.6	54.2	15.6	56.4	17.8
48.1	66.8	18.7	70.0	21.9
51.8	70.8	19.0	72.0	20.2
59.4	73.2	13.8	77.8	18.4
66.7	84.8	18.1	86.6	19.9
C. Size of Cards = 6″ × 9″. N = 100 Subjects				
40.5	57.7	17.2	58.9	18.4
47.6	64.7	17.1	67.4	19.8
53.5	72.4	18.9	74.2	20.7
58.8	79.3	20.5	77.8	19.0
71.6	88.1	16.5	87.6	16.0

center area to total card area in all three card sizes for black on white and for white on black. Minor differences in trend may be noted but these are relatively unimportant. On the average, the overestimation is in the neighborhood of 18 per cent. This is not much different from the 25 per cent overestimation provided by the questionnaire study. In other words, it is apparent that readers tend to believe that the printed area of a book page occupies a larger proportion of the total page space than measurement reveals because of the existence of a part-whole proportion illusion.

SUMMARY

1. Printing rules and printing practice utilize only 50 per cent of available page space for the printed text.

2. Readers are unaware of this fact, believing that the printed text occupies on the average 75 per cent of the page.

3. The discrepancy between belief and fact is due to an experimentally demonstrable part-whole proportion illusion.

4. So far as the writers are aware this illusion has not been described heretofore in the literature.

THE POINT-TALLY

A MODIFIED METHOD OF SCORING THE STRONG'S VOCATIONAL INTEREST BLANK

JOSEPH S. KOPAS
Penn College

INTRODUCTION

IN his most recent book entitled *Aptitudes and Aptitude Testing,* Dr. Walter Bingham makes the following statement concerning the Interest Blank developed by Dr. Edward Strong Jr.: "The Vocational Interest Blank is one of the most valued aids to counseling which prolonged scientific research has produced." Any counselor who has used the test will agree with that statement. The Interest Blank is very widely accepted. However, its use has been somewhat limited, particularly as a group testing device, because of the relatively high cost of scoring it.

The regular method of scoring the Interest Blank by the use of the standard keys is so tedious, difficult, and time consuming—the process takes an experienced person about four and one-half hours—that it is rarely scored in that manner. The user of the test is practically compelled to have it scored by the Hollerith machine. The most serious drawback of that method is its cost which in quantity lots is usually seventy cents per test—not to mention under normal conditions a seven-to-ten-day delay between the time the test is filled out, sent away, scored, and returned.

If scored for seven professions typical of the seven areas into which the twenty-nine professions may be appropriately classified, the cost is reduced to thirty-five cents *per test,* and the scoring time, if done by hand, to approximately one and one-half hours. The disadvantages of this procedure are still

rather important and much the same as previously mentioned when the test is scored in the usual manner. The cost is reduced, but so is the value of the results. Some counselors use the test by merely inspecting the answers to the items. In this way they try to interpret the possible findings of the test. Naturally, the results are rather subjective, incomplete, and unreliable, although helpful.

At Fenn College we have used the Interest Blank as a part of the entrance testing program since 1930. The cost of scoring the test has been approximately three times as great as the cost of the rest of the testing program. The curtailment of expenses during the depression precipitated the problem of either giving up the test, which we were reluctant to do because of its value to our guidance work, or finding some less expensive method of scoring it.

After a careful analysis of the way the test results were used, it was found that from an educational and vocational guidance standpoint the rating on each profession was helpful but not altogether necessary if clearly defined areas of interest were located. In fact, a majority of the counselors were grouping the results of individual professions that correlated highly with each other into patterns or areas of interest, and counseling the testee on that basis. Usually only the first three highest areas of vocational interest were utilized.

Dr. Strong found that the rating on certain professions correlated highly with each other. Therefore grouping professions on that basis is not only logical but is borne out by Dr. Strong's research.[1]

In counseling it is desirable that an individual be directed to test out his vocational and educational plans experimentally in programs available for that purpose. Knowledge of a person's highest areas of vocational interest helps to plan such programs. The additional information that becomes available to a normally functioning guidance department during

[1] Strong, Jr., Edward K. "The Vocational Interest Test." *Occupations*, Vol. XII (April, 1934), 49–56.

the progress of the individual helps to guide him later into more specialized channels.

To be more specific, let us examine the scores obtained by a high school senior on the Interest Blank as shown in Chart A.

Chart A

I		II		III	
A. Physical Science		B. Social Science		C. Business Science	
Key No. 1 Engineering and Technical Occupations		Key No. 5 Literary and Legal Occupations		Key No. 8 Sales Occupations	
Engineer	A	Lawyer	C	Life Ins. Salesman	C
Architect	B−	Journalist	C	Real Estate Salesman	C
Mathematician	B+	Advertiser	C	Key No. 9 Office Occupations	
Physicist	B	Key No. 6 Teaching Occupations		Purchasing Agent	C
Key No. 2. Medical and Biological Occupations		Teacher	C	Clerk	C
Chemist	B+	Minister	C	Accountant	C
Doctor	B−	Key No. 7 Administration		Key No. 10	
Dentist	B−	Personnel Manager	C	Certified Public Accountant	B−
Key No. 3		School Sup't.	C		
Psychologist	B	Physical Director	C		
Key No. 4		Y.M.C.A. Sec'y	C		
Farmer	B				

CHART A: A sample of the results obtained by a high school senior arranged according to the three major divisions and the ten areas of vocational interest.

The highest area of interest is represented by Group 1, which is composed of the following professions: mathematician, engineer, architect, and physicist. This information to a counselor is significant, for it indicates that the high school senior, as far as interest is concerned, would probably do more satisfactory work in an engineering college than in a business administration or liberal arts college. The fact that the individual in question ranks higher on the engineering scale (A) than on the mathematician scale (B+) may be worthwhile information, but of secondary importance. The education program for the first few years of college is, in most cases, the same for the students training for either of the two professions. By the end of the second year the student has had a chance to gain more experience, and the counselor, from the

information gathered on the cumulative record and from personal contact and observation, has had an opportunity to obtain a more complete picture of the student's ability, attitude, interest, and work habits. He therefore does not now need to depend as much on the results of the interest test in directing the student either into engineering or into the field of mathematics.

Probably a good way of summarizing all this would be to say that in most cases knowledge of the highest areas of interest is sufficient and probably as helpful in the early stages of guidance work as the scores on the individual professions.

Various methods of scoring were tried which might reliably identify areas or patterns of vocational interests and would be simple and not too tedious or time consuming. Of the methods tried the point-tally combination came nearest to satisfying the conditions above. Its use experimentally during the past three years in colleges, high schools, Y.M.C.A.'s and community guidance centers (seventeen in total) has demonstrated that the method is economical, practical, and highly satisfactory.

EXPLANATION OF THE DEVELOPMENT OF THE POINT-TALLY METHOD OF SCORING

Twenty-four of the twenty-nine professions for which the test is ordinarily scored were selected for use in this modified method of scoring. The professions which correlated highly with each other were grouped. This process resulted in ten groups. Each group represented an area of vocational interest and each profession was considered a typical representative of the group. In three cases the pattern of vocational interests is represented by only one profession because that particular profession could not be combined logically with any other. Chart A illustrates the grouping of the twenty-four professions.

The highest weighted items of each of the professions constituting a pattern or area of vocational interest were selected

and used in the construction of the scoring keys. These items are represented by points on the scoring keys. When totaled they give the (C) or common score value.

In the process of selecting the highest weighted items, it was discovered that in each area or pattern of interests some of the items ranging from ten to twenty in number not only had a high common value for each profession that constituted the group, but also had a much higher value for one particular profession of the group. These items are indicated on the scoring keys by appropriate letters. (E for engineer, A for architect, etc.) In the process of scoring, these weighted items are tallied and constitute the (T) or tally score value. The final score is the sum of the common score (C) and the tally score (T) values. The maximum score on each of the keys is 100. The "C" and "T" values for each profession are given in Section A of Table I.

Keys No. 3, 4, and 10 cover only one profession. Therefore they have only the points on them and no letters. The number of points used in the scoring procedure on the other keys varies from 80 on Key No. 2 to 95 on Key No. 8. The difference between the number of points on the key and the maximum score of 100 is made up by the additional weighted items represented by the letters on the various scoring keys.

Two kinds of keys were developed to score the Interest Blank, one to be used when the blank is filled out in the regular manner, and the other when a special score sheet is used. The first type mentioned will be identified in this article as Regular Point-Tally Keys and the second as Special Point-Tally Keys.

PROCEDURE FOR SCORING THE INTEREST BLANK BY THE POINT-TALLY METHOD

The scoring keys are made of transparent celluloid and are so designed that any section of the key may be quickly aligned with the respective section of the interest test. After the key and the test are lined up properly for scoring, the points or dots on the celluloid key that coincide with the marked items

on the Interest Blank are counted. Then the extra weighted items represented on the key by letters (A for architect, E for engineer, etc.) are tallied. The two sums are then combined to give the final score. This process is repeated with each of

TABLE I

Composition of Tally Points, Correlation of the Point-Tally Method of Scoring with the Standard Method and Letter Equivalents

		SECTION A			SECTION B	SECTION C				
		C	T	Total	Correlation	Y A	U+ B+	U B	U– B–	N or C
Key 1	Engineer	90	10	100	.60	54	48	43	38	
	Architect	90	10	100	.57	59	53	49	44	
	Mathematician	90	10	100	.63	60	54	50	43	
	Physicist	90	10	100	.70	61	56	51	43	
Key 2	Chemist	80	20	100	.63	56	50	44	38	
	Doctor	80	20	100	.60	56	45	43	38	
	Dentist	80	20	100	.62	50	44	42	37	
Key 3	Psychologist	100	0	100	.71	51	48	41	38	
Key 4	Farmer	100	0	100	.56	56	50	45	39	
Key 5	Lawyer	90	10	100	.64	53	45	37	32	
	Journalist	90	10	100	.70	54	40	36	31	
	Advertiser	90	10	100	.63	55	49	43	36	
Key 6	Teacher	85	15	100	.62	53	45	39	35	
	Minister	85	15	100	.65	51	48	46	39	
Key 7	Personnel Manager	85	15	100	.57	59	49	40	35	
	School Sup't.	85	15	100	.56	61	56	48	44	
	Phys. Director	85	15	100	.56	54	48	41	37	
	Y.M.C.A. Sec'y	85	15	100	.54	57	49	43	38	
Key 8	Life Ins. Salesman	95	05	100	.66	53	48	43	37	
	Real Estate Salesman	95	05	100	.71	56	46	41	36	
Key 9	Purchasing Agent	86	14	100	.66	55	46	38	33	
	Clerical Worker	86	14	100	.67	56	46	40	36	
	Accountant	86	14	100	.66	62	55	45	40	
Key 10	Certified Public Accountant	100	0	100	.49	53	45	38	32	

Everything below the lower limit of U– (B–)

the other nine keys. The letter equivalent for any of the twenty-four professions is obtained from the data in Section C of Table I.

THE ADVANTAGES AND DISADVANTAGES OF THE POINT-TALLY METHODS OF SCORING THE INTEREST BLANK

The major advantages that come from the use of the Point-Tally method of scoring the Interest Blank as compared to the method worked out by Dr. Strong are those of economy of cost and time, and the simplicity. With the use of the Point-Tally Keys the Interest Blank can be scored in half an hour or less by almost any person. No special skill or training is necessary.

As can be seen from the data in Table II, this modified method of scoring is highly effective in identifying the highest areas of vocational interests. The results obtained from this method on individual professions also correlate fairly well with the results obtained by the standard method of scoring. The correlations for each profession are listed in Section B of Table I. They range from .49 to .71. These results are rather interesting, for this method of scoring was designed to spot the highest patterns of vocational interests and not for the purpose of measuring results of individual professions.

The Special Point-Tally Keys as compared with the Regular Point-Tally Keys are smaller, easier to handle, and need to be aligned only once for the entire test. The Regular Point-Tally Keys have to be aligned for each of the eight sections of the Vocational Interest Blank and the pages have to be handled in much the same manner as is the case when the test is originally filled out.

When the Score Sheet is used, the test data is all on one sheet. For this reason the scoring time is reduced to approximately twenty minutes, an additional saving of ten minutes. Filing space is also conserved, for the Score Sheet is naturally less bulky to file than the Interest Blank itself. The Interest Blank can be used over and over again, because the individual taking the test does not write anything on the blank. Since

the cost of a Score Sheet is considerably less than the cost of the Interest Blank, a rather significant saving results through the use of it. The taking of the test is a little more complicated when the Point-Tally Score Sheet is used and therefore more attention must be exercised by the examiner and more explanation given. The testee must exercise greater care to avoid misalignment or the mistaking of sections. It is also necessary to mark in a special manner the Interest Blanks.

The chief advantages of the two Point-Tally Methods of scoring the Interest Blank may be briefly summarized as follows: The Regular Point-Tally Keys make possible the use of the test in a normal way, cut the scoring time to approximately thirty minutes, simplify the scoring very much, and spot the highest areas of interest reliably.

The Special Point-Tally Keys, in addition, make it possible to use the Interest Blank over and over again, and to further reduce the time of scoring by approximately ten minutes.

COMPARATIVE RESULTS

In comparing results of the Point-Tally Method of scoring the Interest Blank with the method worked out by Dr. Strong, 313 tests were scored both ways. This number in some cases has been insufficient to get highly reliable results. More research is necessary. Further difficulty was encountered in trying to find some method of correlating the results on the basis of three highest areas of interest, since the average counselor considers only the first, second, and third highest areas of vocational interest in counseling, and the keys were designed with that purpose in mind.

No method has been found by the author for correlating the results of the three highest areas as measured by the Point-Tally Method with the three highest areas out of the ten possibilities obtained by the standard method of scoring. Therefore the comparisons were made in terms of percentages.

Table II shows the number of cases and gives the percentage of the times the first, second, and third highest areas of voca-

tional interest obtained by the Point-Tally Method checked with the first, second, and third highest areas of the standard method of scoring.

TABLE II

Comparison of the Results with the Standard Method of Scoring on the Basis of the Three Highest of Ten Areas of Interest

RANK	FIRST		SECOND		THIRD		OUTSIDE	
	No. of Cases	%	No. of Cases	%	No. of Cases	%	No. of Cases	%
First	256	82	44	14	7	2	6	2
Second	37	12	176	56	57	18	43	14
Third	7	2	61	19	151	48	94	30

In eighty-two per cent of the cases the highest areas of vocational interest were found to be the same, the second highest checked identically in fifty-six per cent of the cases, and the third highest in forty-eight per cent of the cases. Ninety-six per cent of the cases that were first in importance in the standard method of scoring fell within the three highest areas obtained by the modified method of scoring. In only two per cent of the cases did the highest areas of interest measured by the Point-Tally Method fall outside of the three highest areas of interest obtained by Dr. Strong's method of scoring. These results are rather encouraging in view of the fact that there are ten categories and the highest choice may range anywhere from first (highest) to tenth (lowest possible rank).

In Table III the comparison was made only on the basis of the three broad categories or divisions of vocational interest—the physical, the social, and the business science divisions. These are represented by the professions shown on Chart A. This kind of a classification is useful in helping high school graduates make a choice between the engineering, the arts and science, and the business administration college programs, and also in selecting work in line with one's interests.

In ninety per cent of the cases the first choices were the same and in seven per cent of the cases the highest area of vocational interest obtained by the standard method of scoring came out

TABLE III

Comparison of the Results on the Basis of Three Major Divisions—(1) Physical, (2) Social, and (3) Business Science Fields

RANK	FIRST		SECOND		THIRD	
	No.	%	No.	%	No.	%
First	282	90	25	8	6	2
Second	22	7	246	78	47	15
Third	9	3	42	14	260	83

second when the Point-Tally Method of scoring was used.

One would naturally expect that the engineering freshmen should rank highest in the physical science field, the business administration freshmen highest in the business science field, and the liberal arts freshmen highest in the social science field. The results for 139 freshmen of Fenn College that took the Interest Blank in the fall of 1936 were classified. Table IV indicates the results.

TABLE IV

Classification According to Fields of Specializations of College Freshmen

	NO. OF CASES	IN THE FIELD						NOT IN THE FIELD			
		Y		U		N		Y		U	
		No.	%	No.	%	No.	%	No.	%	No.	%
Engineering Freshmen	86	47	55	33	38	6	7	11	13	52	60
Bus. Administration Freshmen	29	17	59	9	31	3	10	5	17	15	51
Liberal Arts Freshmen	24	13	54	7	29	4	17	5	21	21	87

Fifty-five per cent of the engineering freshmen received scores of Y (yes), thirty-eight per cent U (uncertain), seven per cent N (no) in the physical science field. To identify the method of scoring and avoid confusion and misunderstanding, the letters Y, U+, U, U–, and N were used in place of A, B+, B, B–, and C respectively. Y, U, and N represent "yes," "uncertain," and "no" respectively and fit in well as general answers to the question, "Does the individual have the interests characteristic of a particular area of interest?" Thirteen

per cent of them received "Y" (yes), and sixty per cent "U" (uncertain) in the other two fields. Of the twenty-nine business administration freshmen, fifty-nine per cent received "U" ranking in the business science field as compared to seventeen per cent that received a "U" ranking in the other two fields. This is a ratio of three to one in favor of the field of specialization.

WHERE THE POINT-TALLY METHOD OF SCORING MAY BE USED

This modified method of scoring the test for the highest areas of interest will be found helpful and useful in colleges, high schools, Y.M.C.A.'s, in community guidance centers, and in almost any situation where group testing is utilized and some method of follow up practiced.

It can also be used as a selective scoring device for identifying the professions most likely to be significant in counseling. These professions can then be scored by the standard Vocational Interest Blank score keys. This system would eliminate the scoring on the professions that normally rate only "C" in value.

BRIEF REPORTS ON MENTAL TESTING

SEX DIFFERENCES IN INTELLIGENCE AMONG NEGRO COLLEGE FRESHMEN

HERMAN G. CANADY

West Virginia State College

Subjects: 1306 Negro college freshmen; 637 males, 669 females, who entered West Virginia State College from 1931 to 1937 inclusive.

TABLE I

Distribution of the Gross Scores on the American Council Psychological Examination of 637 Males and 669 Females, the Percents of Each Sex Attaining the Various Scores, the Ranges, Means, Standard Deviations and Pearson's Coefficients of Variation for Each Group

SCORES	MALE		FEMALE	
	No.	Percent	No.	Percent
230–244	2	.31	1	.15
215–229	3	.47	1	.15
200–214	6	.94	4	.60
185–199	5	.79	7	1.05
170–184	15	2.36	13	2.03
155–169	20	3.14	12	1.78
140–154	17	2.67	24	3.59
125–139	45	7.06	37	5.53
110–124	40	6.28	53	7.93
95–109	68	10.68	69	10.33
80– 94	76	11.93	85	12.70
65– 79	88	13.82	101	15.05
50– 64	109	17.11	114	17.00
35– 49	81	12.72	71	10.60
20– 34	45	7.06	56	8.36
5– 19	17	2.66	21	3.14
Total	637	100.00	669	100.00
Highest Score	235		231	
Lowest Score	5		7	
Range	230		224	
Median	76.34		75.77	
Mean	84.50		81.50	
S.D.	43.36		41.40	
Pearson's c. of v.	51.30		50.80	

Test: American Council on Education Psychological Examination.

Results: (Tables on attached pages).

Conclusions: The following conclusions are limited to this particular group of students, material and conditions of investigation:

1. No significant sex difference in "general intelligence" (gross scores) was discovered. In all comparisons of differences in central tendency and variability, the similarity rather than the difference was the striking phenomenon.

This finding is not in agreement with that of Caliver. In his second investigation (1933), he found a significant and reliable difference in "lump scores" in favor of the males.

Our group, it appears, conforms more to the general picture as revealed by the large majority of studies among white students in showing: (1) a tremendous amount of overlapping between the two groups; and (2) that the differences in variability in favor of either sex are too small to be

TABLE II

Means, Sigmas of the Distributions, and Sigmas of the Means of Males (637) and Females (669) for the Five Sub-Test Scores and Gross Scores on the American Council Psychological Examination

TEST	MEASURE	MALE	FEMALE
Completion	Mean	19.40	17.56
	σ (dis.)	11.00	10.05
	σ (mean)	.43	.39
Artificial Language	Mean	16.25	19.20
	σ (dis.)	11.50	12.95
	σ (mean)	.45	.50
Analogies	Mean	14.30	14.40
	σ (dis.)	9.57	10.57
	σ (mean)	.37	.41
Arithmetic	Mean	13.70	10.85
	σ (dis.)	10.58	9.00
	σ (mean)	.43	.35
Opposites	Mean	22.65	21.85
	σ (dis.)	16.68	16.61
	σ (mean)	.65	.64
GROSS SCORE	Mean	84.50	81.50
	σ (dis.)	43.36	41.40
	σ (mean)	1.72	1.60

of much significance. The results therefore, do not favor the theory of a greater variability of the male sex.

2. The analysis of the sub-tests revealed a large and reliable sex difference in performance on the verbal or linguistic and numerical parts of the test, the latter in favor of the males and the former in favor of the females. Identical results among white students have been discovered by practically all investigators.

Caliver found (1933) that in every case except one, artificial language, the males surpassed the females. Our comparisons, however, show differences in favor of the females in three instances—artificial language, analogies, and opposites.

3. Finally, our data tend to support the growing tendency to look for sex differences in specific traits or separate abilities rather than in "general level of performance."

TABLE III

Sex Comparisons in Terms of the Chances of a True Difference

TESTS	DIFFERENCES BETWEEN MEANS D	σ (DIFF.)	CRITICAL RATIO $\frac{D}{\sigma \text{ (DIFF.)}}$	CHANCES IN ONE HUNDRED THAT THE TRUE DIFFERENCE IS GREATER THAN ZERO[1]	DIFFERENCES IN FAVOR OF
Completion	1.84	.487	3.79	100	Male
Artificial Language	2.95	.673	4.30	100	Female
Analogies	.10	.552	.18	58	Female
Arithmetic	2.80	.455	6.20	100	Male
Opposites	.80	.913	.87	80	Female
Gross Scores	3.00	2.350	1.28	90	Male

[1] Garrett, Henry E., *Statistics in Psychology and Education.* Table XIV, pp. 134.

REFERENCES

HERMAN G. CANADY. "Individual Differences among Freshmen at West Virginia State College." *Journal of Negro Education*, 5, pp. 246–58, April, 1935.

AMBROSE CALIVER. *A Personnel Study of Negro College Students*, Teachers College Contributions to Education, No. 484, 1931, pp. 67–70.

AMBROSE CALIVER. *A Background Study of Negro College Students*, U. S. Office of Education Bulletin No. 8, 1933, pp. 13–18.

FRANK S. FREEMAN. *Individual Differences*, 1934, Chap. 6.

NEWS AND NOTES

Professor F. Kenneth Berrien, formerly of the Department of Psychology, University of Pennsylvania, has recently joined the staff of the Department of Psychology at Colgate University. Fred S. Keller has resigned at Colgate to join the staff of Columbia University.

The Seventh Educational Conference will be sponsored by the Committee on Measurement and Guidance of the American Council on Education, the Commission on the Relation of School and College of the Progressive Education Association, the Cooperative Test Service, and the Educational Records Bureau. The meetings will be held at the Hotel Roosevelt in New York City, October 27 and 28 and will be open to all educators, whether or not they hold membership in the sponsoring organizations. Dr. Karl W. Bigelow, Director of the Commission on Teacher Education of the American Council on Education, will speak at the opening session on the crucial problem of the training and selection of teachers. Dr. George D. Stoddard, University of Iowa, will present an analysis and interpretation of data which indicate that the average IQ of about two hundred children was increased about twenty points in a period of three or four years.

Dr. Samuel T. Orton will present a paper on the nature and cure of reading difficulties. At the annual dinner on Thursday evening Dr. Francis T. Spaulding, Harvard University, and Regent John Lord O'Brian, of New York State, will present some of the results of the New York State Regents' Inquiry into the cost and character of education. The importance of the library in education will be discussed by Dr. B. Lamar Johnson, Stephens College, and followed by a paper by Dr. Robert M. Bear on the very significant remedial reading and eye clinic work that has been going on at Dartmouth for some years. The speaker at the Friday luncheon session will be President H. W. Dodds, of Princeton.

The growth of state-wide testing programs in several states has been one of the most significant features of American education during the last decade. On Friday afternoon Dr. Walter V. Bingham will speak on "A National Perspective on Testing and Guidance." Some workers in the field, impressed by the dangers of unnecessary duplication of services and by the need for more effective research, feel that the problem of coordination is one of the most important that confronts us at this time. Dr. V. A. C. Henmon, University of Wisconsin, will present an address on "The

Scope and Character of State Testing Programs,'' followed by a paper on ''The Use of Test Results'' by Dr. H. T. Manuel, University of Texas.

Those who are planning to attend the Conference should make reservations in advance through the Educational Records Bureau, 437 West 59th Street, New York City.

The American Nautical Academy, National Training School for Merchant Marine Officers, Washington, D. C., is offering the ninth annual summer course for boys and young men between the ages of 10 and 25 years to secure practical ship experience on board a training ship from July 1 to October 1. The only required expense is for meals. The boys will learn the regular daily ship routine and will be given instruction in nautical subjects.

Those completing the summer course with a passing grade will be eligible to apply for a scholarship in the Regular Course. The latter covers a period of three years. The first two years the cadet spends cruising and studying on the training ship. For the third year he is assigned to a steamship of one of the merchant marine lines. The successful cadet is now eligible to take the officer's examination of the United States Bureau of Navigation and Inspection, Department of Commerce, for a certificate as a Third Officer in the Deck Division of steamships in the Merchant Service of the United States. Further information concerning these courses may be obtained from the American Nautical Academy, National Training School for Merchant Marine Officers, Washington, D. C.

The Administration Building, Ohio State University, Columbus, Ohio, will be the headquarters for the meetings of the American Association of Applied Psychology, from September 5 to 8 inclusive. To indicate the scope of the Association and how already it has attempted to make preparations for its future work, we list the following group meetings in full.

Group A. Committees on Standards of Training.
W. V. Bingham, Coordinating Chairman
Faculty Club, South Lounge

1. Committee on Applied Instruction in College and Secondary Schools. Edwin R. Henry, Chairman
2. Committee on Dissertations in Applied Psychology. A. T. Poffenberger, Chairman
3. Committee on Graduate Instruction in Applied Psychology. Frank N. Freeman, Chairman

Group B. Committees on Internal Activities of the Profession.
Douglas H. Fryer, Coordinating Chairman
Faculty Club, Grand Lounge

1. Committee on Applied Journals
   James P. Porter, Chairman
2. Committee on Professional Employment
   Robert Hoppock, Chairman
3. Committee on Professional Ethics
   Warren W. Coxe, Chairman
4. Committee on Working Conditions of Applied Psychology
   E. Lowell Kelly, Chairman

Group C. Committees on Applied Techniques.
Robert T. Rock, Coordinating Chairman
Derby Hall, Room 202

1. Committee on Standardized Tests
   Henry E. Garrett, Chairman
2. Committee on Technical Manuals
   Herbert A. Toops, Chairman
3. Committee on Standardized Test-Norms
   Jack W. Dunlap, Chairman

Group D. Committees on Public and Professional Relations.
John G. Jenkins, Coordinating Chairman
Derby Hall, Room 100

1. Committee on Relations with Educational Profession
   M. R. Trabue, Chairman
2. Committee on Relations with Engineering Profession
   Henry C. Link, Chairman
3. Committee on Relations with Medical Profession
   Francis N. Maxfield, Chairman
4. Committee on Relations with Social Workers' Profession
   Edwina A. Cowan, Chairman
5. Committee on Public Exhibits
   Richard H. Paynter, Chairman
6. Committee on Relations with Government
   Edmund G. Williamson, Chairman

These meetings of the professional committees will be held from 10: 00 to 12: 00 on Monday, September 5. A full and interesting program has been provided for the four-day meetings. Of greatest interest to the readers of THE JOURNAL OF APPLIED PSYCHOLOGY we believe will be the mention of the following symposia: "The Work of Psychologists in Institutions and Child Guidance Clinics," 9: 00–10: 00, Tuesday, September 6, Derby Hall, Room 100, Andrew W. Brown, Chairman; "The Training and Duties of the School Psychologist," at 10: 30 to 12: 00, Tuesday, Auditorium, Social Administration Building, with Gertrude Hildreth as Chairman; "What Can Psychology Contribute to Safety of the Highways," from 4: 30 to 6: 00, Tuesday, Alvhh R. Lauer, Chairman, Social Administration Building, Auditorium.

The Presidential Dinner will be held at 6:00 Wednesday, September 7, at the Faculty Club. The address of welcome will be given by Dr. G. F. Arps and the presidential address on "Applied and Professional Attitudes" by Dr. Douglas H. Fryer. Thursday afternoon will be devoted to research programs, with Dean A. Worcester as Chairman, held in the Chemistry Building, Room 100. Among those appearing on this program are H. H. Remmers, Richard W. Husband, A. R. Lauer, and Victor H. Noll. Dr. Herbert A. Toops, Ohio State University, is chairman of the Program Committee.

The American Psychological Association will hold its annual meetings at Ohio State University, Columbus, Ohio, from September 7 to 10 inclusive. Information concerning the program may be obtained from the Secretary of the Association, Dr. Willard Olson, University of Michigan, Ann Arbor, Michigan.

Rio de Janeiro will be host to the Eighth Biennial Conference of the World Federation of Education Associations during the summer of 1939. Dr. Paul Monroe, President of the Associations believes that this meeting will further considerably the development of the Good Neighbor Policy between our Government and the Latin American Republics and also provide an excellent opportunity to teachers of the United States and Canada to visit South America under most desirable auspices. Special arrangements are being made to keep the cost of the journey as low as possible. Large transatlantic liners are being chartered and will follow the most interesting and varied itineraries, including visits to many South American ports and to the more important West Indies Islands. Two cruises each of approximately 57 days will leave New York about July 1. The minimum fare will be about $500. Literature and full information regarding these travel plans will be available through the W.F.E.A. Travel Bureau, National Education Association Building, 1201 Sixteenth Street, N.W., Washington, D. C., and also through authorized travel agents throughout the United States and Canada.

The Seventh International Management Congress will open its five-day series of General and Technical Sessions in Washington, D. C., on September 19, continuing through September 23. The two official themes of the Congress will be "Recent Developments in Management" and "Economic and Social Aspects of Management," developed by experts in the field of management. More than two hundred papers on a wide range of subjects relating to Administration, Production, Distribution, Personnel, Agriculture, and the Home have been printed in six volumes and will be distributed to all registrants prior to the Congress. These papers will

serve as a basis for discussion at the Technical Sessions. Membership in the Congress is open to anyone interested in the problems of management upon the payment of a registration fee of ten dollars. Further information may be obtained from Nathaniel W. Barnes, Executive Secretary, 347 Madison Avenue, Room 1201, New York City.

The H. W. Wilson Company, 950 University Avenue, New York, announces the publication of the second revised edition of INDEX TO VOCATIONS, by Willodeen Price and Zelma E. Ticen. The new edition adds eight new books and 230 monographs to the 115 books analysed in the first edition, and contains an alphabetical arrangement of the Index by name of the specific vocation; the author, title, date and inclusive paging given for each entry; the biographical vocational entries under a separate listing; and a useful selected list of recent books for vocational teachers and counselors.

BOOK REVIEWS

HARRINGTON, MILTON. *A Biological Approach to the Problem of Abnormal Behavior.* (Copyrighted 1938 by Milton Harrington.) Distributed by The Science Press Printing Company, Lancaster, Pa. Pp. 459.

In contrast to the animistic, which he also at times calls motivistic and psychological point of view of Freud, Adler, and their ilk, the writer gives us a mechanistic, physiological, or biological approach. The book is divided into three parts. Part One, bearing the sub-title "First Principles," begins with a clear but elementary discussion of basic concepts and methods, treats the problems of free-will and of consciousness, and closes with an evaluation of the "old psychology." This old psychology is that of the psychoanalysts.

Part Two, called "Psycho-physiology," is simply an outline of several of the main topics of general psychology discussed in terms of physiological mechanisms. Perhaps as fair an example as any of the treatment of physiological psychology may be taken from one of the early chapters in this section entitled "Learning and the Law of Confluence." After likening what goes on in the nervous system when we learn to the deepening of stream beds in a sandy plain, the writer raises the question of why one path rather than another comes to predominate. He says, "It would seem that when the organism is subjected to two or more stimuli at one and the same time, there is a tendency for the streams of energy which they release to flow together, for the impulses to coalesce; so that where we have what may be looked upon as two separate and distinct impulses entering the nervous system upon its afferent side from two different sense organs or sets of sense organs, we have only one impulse or set of impulses emerging on the efferent side to give rise to action of muscles and glands. This tendency for two or more impulses passing through the nervous system at the same time to flow together is . . . an extremely important one. Because it is of such great importance it is necessary to give it a name. We shall here call it the Law of Confluence." (p. 84.)

Or again, in the physiological explanation of memory the author says, "Memory recall, therefore, we may suppose, is simply a matter of sending streams of nervous energy flowing over the same combination of paths as that over which the impulses flowed which gave rise to the original image." (p. 116.)

Having cleared the way the author comes in Part Three to his psycho-

pathology. Abnormal behavior is explained in terms of the bio-mechanistic view already developed. Defects in the nerve cells, defects in the organization of the nervous system, and wrong learning become the main physiological reasons for abnormal behavior. The nature of the stimulus, in the form of what might be called impossible situations, is also taken into account as a reason for such behavior. Since these factors are determined by heredity, training, health and environment, mental health becomes a matter of dealing adequately with the latter four factors.

The author is attempting to develop a theory, and states that since this is the case, the book should be regarded as a bare outline any chapter of which could be expanded into a volume. Considered as an outline it is excellently done, being closely integrated and logically organized.

Considered, however, as a theoretical contribution, one may raise the question of whether its significance is all that the writer seems to believe it to be. Psychology is to the writer, as far as he betrays in his book, the psychology of Freud. Were it not for a few scattered references to such writers as James and Jastrow one would almost think that the writer is not aware of the existence of the psychology of the psychologists, and had created, *de novo*, a pretty orthodox psychology of his own. The student of psychology will recognize the physiological explanation of learning as essentially what he in all probability learned from his beginning text, with the addition of the law of confluence which will be seen to be either the drainage theory or just about the same thing as the drainage theory. The physiological explanation of emotion is, of course, as the author states at the end of the chapter, the James-Lange theory, and so on throughout. Certainly physiological explanations, such as are presented in this section, the psychophysiology, are current coin to even the undergraduate major in psychology.

It is, of course, in the third division of his book that the writer seeks to make his real contribution; parts one and two being only introductory to this, the real subject of discussion. If the point of view here presented were essentially new to the psychiatrists it would either be revolutionary or, more likely, fail to reach them at all. Needless to say, it is not new, the writer himself recognizing that Dr. Adolf Meyer's view is a biological one, although Dr. Meyer does not follow his theory through to its logical conclusions, according to the writer. While not questioning the originality of many of the interpretations of the writer, certainly both the biological point of view and many of the interpretations given are sufficiently familiar so that the ground is well prepared for the consistently biological foundation presented here.

The significance of the book lies, then, not in any new factual material, not primarily in new psychological concepts or explanations, but in the fact that it attempts to be consistently biological, presenting as its title states a biological approach to the problem of abnormal behavior.

The practical application of the principles developed in this volume, which the author says will take us into the field of what he calls psychoorthology, is reserved for a subsequent volume.

AMOS C. ANDERSON,
Ohio University.

O. B. DOUGLAS AND B. F. HOLLAND. *Fundamentals of Educational Psychology.* New York, The MacMillan Company, 1938. xv + 598. $2.50.

This book is a text designed for use in a first or elementary course in educational psychology. The authors report that they have satisfactorily used the material in a mimeographed edition at the University of Texas. College sophomores, then, will be the most frequent readers in beginning their study of psychology. It is important to keep this in mind in evaluating this text.

This book is unique in several respects. In the first place each chapter is so sub-divided that one receives an impression of a detailed outline. In the second place each chapter begins with a section explaining the purpose of the chapter. Finally the authors have attempted to define systematically all psychological terms. As a result the text teems with definitions. Indeed, it might appear to a beginner that psychology is nothing more than a special vocabulary. Were it not for the fact that this book has been found teachable by the authors, an instructor faced with a class of college sophomores might think the book unattractive.

The material is divided into four parts. Part I serves to introduce the student to the content of educational psychology. It makes a useful contribution by including a good orientation to the method of science in general the relation of psychology to education. The question arises, however, whether this scientific orientation belongs in the province of educational psychology.

Part II includes a discussion of environment and heredity. Considerable space is devoted to the sense organs, muscles, and nervous system. The authors deserve a word of merit for their clear and simple figures. In this treatment they present a novel viewpoint of the brain as a reacting mechanism just as the muscles and glands are reactors. The brain is also traditionally considered as a part of the connection and integrating system.

Learning is discussed in part III. The authors accept the synapse theory in describing learning and in further treatment of the numerous problems. Each chapter is a description of a type of learning. The separate presentation of these types is quite unusual. These types of learning include Sensori-Motor (simple motor habits), Perceptual-Motor (social habits, complex motor skills), Perceptual (language meaning),

Conceptual (general ideas, principles), Associative (ideas, verbatim information), and Imaginative (understanding of things not present). Each type is described in detail including its definition, its occurrence, its nature, and its relation to school subjects. The problems of personality and social and personal adjustment are found in the chapters on Perceptual Motor Learning.

In explaining all types of learning, we authors frequently fall back upon a principle of contiguous stimulation. It is almost a unifying principle that binds together the different types of learning. The law states that when two or more stimuli are presented at the same time, a connection is formed between each stimulus and the total responses to these stimuli. Therefore, the subsequent occurrence of any one stimulus tends to evoke any or all of the responses in the total reaction. Thus, the law is in actuality a re-statement of the conditioned response.

The effort to relate the types of learning to school subjects is a good contribution. For example, in the discussion of Imaginative Learning the suggestion is made that the usual chronological teaching of history could be profitably supplemented by beginning with the present and working to the past.

Tests and Measurement are presented in part IV. There is a valuable inclusion of physical and sensory tests that are usually slighted. The treatment of intelligence tests, achievement tests, and tests of other aptitudes is somewhat brief. The final chapter of the book is a treatment of statistics in educational measurement. This presentation is exceptionally fine.

This book is recommended for use where the instructor supplements the text material with discussions and readings of the experimental work in the field of educational psychology. Such supplementation becomes necessary because the text only briefly mentions these original experiments and seldom refers to the sources.

SIDNEY ROSLOW,
Psychological Corporation.

Social Life and Personality, by E. S. BOGARDUS and R. H. LEWIS. New York: Silver Burdett & Co., 1938. Pp. 581.

While this is a book written for high school students, it has many suggestions in it which might well be taken by college instructors. I believe that in technique of teaching, the colleges today lag considerably behind the secondary schools. The use of questions, study outlines, activities, etc., in this book might well be imitated by more "psychological" volumes.

The term "personality" in the title is somewhat misleading. The book is concerned in the main with social conditions in modern American

life—family, play, education, industry, national relations, religion and art. The connection between these large-scale social situations and the personality of individual workers, while very real to a psychologist, is not necessarily clear to high school students. I do not believe that the authors have done an adequate job of showing the student the relationship between the individual and his group. For example, a picture shows a skilled silver engraver and a worker at a machine lathe. The question is asked, "How does the work of these men differ in the effect on personality?" This is a good question, but inadequate material for generalizing about the effects of occupation is given in the text. I should have preferred to see one such example worked out for the reader, followed by other questions for him to answer.

I also deplore the absence of some material on the forms of behavior of the individual—self-expressive reactions, rationalization, repression, compensation, etc. I think it is important for high school students to be informed about such behavior.

On the other hand, one must grant that the book is very stimulating and well-written. The treatment of specific topics which often suffer in public school texts, *e.g.*, labor unions, is fair and reasonably clear. The treatment of the problem of war is perhaps excessively psychological (strange words for a psychologist!). I think a little more emphasis on trade and investments would be appropriate.

If more high school students worked with texts like this, college instructors would complain less about the inadequacy of their students' preparation.

ROSS STAGNER,
University of Akron.

HOWE, E. GRAHAM. *War Dance: a Study of the Psychology of War.* London: Faber & Faber Ltd., 1937. Pp. 315.

Despite the tremendous importance of war as a psychological problem, it has been given few systematic discussions by academically trained psychologists. In view of this fact, more attention must be given to this curious little book than it might otherwise merit.

The author states that his book is a translation of subject matter given in a series of six lectures to the Home and School Council. The reviewer wonders what the unfortunate listeners heard, for even in print the style is so abstruse and disjoined as to require frequent re-reading to locate a connection.

Despite these stylistic defects, the book definitely has something to offer. We find, for example, a very reasonable remark in the very beginning: "If there are 'Mad Dogs' at large in Europe at the present time (but how dangerous are these snappy labels!) it is not enough to ask for guns or chains. Who bit the dog and why is he so mad?" (P. 39.)

It is the author's thesis that men do not want war; that wars are due to misunderstandings. One country's people distrust and fear those of alien nationality; and across the frontier the same sentiments are projected back. But, points out Dr. Howe, "Even in our search for peace there are so many of us who do not help to ease the strain of misunderstanding. They are the ones who would, with reasonable gesture and excellent intention, either eliminate the problem altogether or *eliminate the one whom they think threatens to be the aggressor.* (Italics the reviewer's.) But how can such aggressiveness ever cure aggressiveness? How can war end war, or intolerance be made to cure intolerance?" (P. 41.) This will not be taken well in the present popularity of "collective security" movements. If civilization survives the next holocaust, some people will have to start practicing this dictum or we shall see the same vicious cycle of war, defeat, despair, hatred, revenge and war going on and ever on.

Everyone wants peace—at least, so we are told. Hence Howe asks, what are the forces operating against peace? "We have to recognize the aggresaive influence in so many places, exercised in so many ways: *especially, however, in the ulterior motives of those politicians who are inclined to solve their own problems vicariously in national disguise,* leading to a degree of earnestness which provokes further misunderstanding." (P. 217—italics the reviewer's.) Among the forces he lists are such as the following: the differentiation of cultures and the recognition of, and insistence upon, this difference; the aggressiveness of desire, even the desire for peace; identification of selfish interests with larger groups; defense against uncertainty and suspense; vested interests and habits; refusal of interdependent groups to change together (others attempt to preserve the *status quo* while one or more change.)

Some of these seem particularly relevant to a psychological interpretation of war. "Desire is usually complicated, however, by ulterior motives, so that we seem to want not for ourselves, but for others. The mother 'unselfishly' wants the advantage of her children, but is this unselfish if she has identified herself with them? . . . The politician is sure the other side is wrong, but what of the projection of his own repressed unrest? The statesman is solely interested in his country's welfare, but what if his country has become confused with his own insatiable egotism?" (P. 218.) And again, "Defence against the uncertainty of suspense fixes advantage and privilege wherever possible by treaty or promise, customs barrier or 'favored nation' clause." "Vested interests are innumerable, and no one who is enjoying them can bear to think that they should be sacrificed for any larger cause. They become moralized: 'This is so, and so this ought to be.'" (P. 219.)

Digging out the author's theory from the debris of dreams, diagrams

and definitions beneath which it is concealed, we may summarize it about as follows: In the world at present are groups of people (nationalities) who are different from one another and acutely conscious of that fact (effect of schools, etc.) Within each of these groups exists a certain pattern of socio-economic relations which provide privilege and power for some, a modicum of security for all. Now it is impossible for such a group to remain static; and yet any attempt on the part of one group to change the existing balance of trade privileges and political power will be met with intense resistance by other groups. Thus each nation fights, not for the sake of war, which is hated and feared, but for the sake of goals which they perceive as attainable only through violence (effect of historical tradition). As this threat of social change increases the insecurity of the population, persons with vested interests to protect and those who are selfishly involved by virtue of identification with national symbols take the lead in expressing hostility to the group which is perceived as the immediate source of this insecurity. "The threatening fist of one induces an equal and opposite gesture of revenge." And so we go to war.

But what of peace? "The way of peace in politics is to act as mediator to these opposing forces which are ranged against one another. *This is not to say that we are sentimentally to deny difference around a conference table,* creating a fog or formula to cover our distress, behind which the forces of evil are more free to bring disaster. . . . Our method must be to define and accept all differences that may be true in fact." (Pp. 228–229.)

While appraising the politicians and their contribution to international unrest through the projection of their own personal dissatisfactions, as we have shown above, the author takes due heed of the fact that the general population is an effective force, and that a different economic system might relieve some of the pressure. "There must be fear, and anxiety is a necessary condition of our living truly in suspense. But there need not be so much, where we can at least be sure of the wherewithal to live, both for ourselves and our families. . . . Relationships within a community would be less aggressive, where all could be assured of the right to at least the necessaries of life." (P. 244.) Here we are reminded of the fact that tensions accumulated as a result of economic and personal insecurity may be projected upon the identifying symbols of another country, and the need for security attained through identification with our own group and especially with our leaders.

As was stated at the beginning of this review, the book is a very curious one. A blend of psychoanalysis and Zen Buddhism (both of which recognize the polarity of psychological reality) is presented through constantly changing metaphors, many of which are only slightly relevant to the topic

apparently under discussion. Odd diagrams decorate and confuse the pages. A number of uninterpreted dreams are incorporated as additional devices for advancing the bepuzzlement of the reader. Only a few of the 300-odd pages deal directly with war.

Despite these serious shortcomings the book is provocative and it includes some ideas of real merit. Psychologists who like puzzles can spend an entertaining evening working through its pages. There is a message hidden there which American psychology needs—it is unfortunate that the author did not state it plainly, but must conceal it in a cryptogram.

Ross Stagner,
University of Akron.

NEW BOOKS AND PAMPHLETS RECEIVED

New books and pamphlets for review should be sent to James P. Porter, Editor, JOURNAL OF APPLIED PSYCHOLOGY, Ohio University, Athens, Ohio.

A Bibliography of Canadian Education. ALBERT H. SMITH and Others. Department of Educational Research, University of Toronto, 371 Bloor St., W., Toronto 5. 1938. $1.00. 302 pp.

General Psychology, Introduction to. JAMES B. STROUD. Prentice-Hall, New York, 1938. $3.25. 675 pp.

My Mind a Kingdom. GEORGE THOMAS. E. P. Dutton, New York, 1938. $2.50. 294 pp.

Plan for Marriage. JOSEPH KIRK FOLSOM. Harper and Brothers, New York, 1938. 305 pp.

Psycho-Analysis. ERNEST JONES. William Wood & Company, Baltimore, 1938. 643 pp.

Psychology in Everday Living. FRANCIS F. POWERS, T. R. MCCONNELL, WILLIAM C. TROW, BRUCE V. MOORE and CHARLES E. SKINNER. D. C. Heath & Company, New York, 1938. 511 pp.

Psychology of Music. CARL E. SEASHORE. McGraw-Hill Company, New York, 1938. $4.00. 397 pp.

Public Opinion and The Individual. GARDNER MURPHY and RENSIS LIKERT. Harper & Brothers, New York, 1938. $3.00. 316 pp.

Social Psychology. DANIEL KATZ and RICHARD L. SCHANCK. John Wiley & Sons, New York, 1938. $3.75. 700 pp.

A Student Personnel Program for Higher Education. ESTHER MCD. LLOYD-JONES and MARGARET RUTH SMITH. McGraw-Hill Company, New York, 1938. $2.75. 322 pp.

The Troubled Mind. C. S. BLUEMEL. Williams and Wilkins Company, Baltimore, 1938. $3.50. 520 pp.

Directory of American Psychological Periodicals

(Continued)

Journal of Comparative Psychology—Baltimore: The Williams & Wilkins Co. Subscription $5.00 per volume of 450 pages. Two volumes every year. Edited by Knight Dunlap and Robert M. Yerkes. Founded 1921.

Comparative Psychology Monographs—Baltimore: Johns Hopkins University Press. Subscription $5.00. 400 pages per volume. Knight Dunlap, Managing Editor. Published without fixed dates, each number a single research. Founded 1922.

Genetic Psychology Monographs—Provincetown, Mass.: The Journal Press. Subscription $7.00 per volume. One volume per year. Edited by Carl Murchison and an international cooperating board. Quarterly. Each number one complete research. Child behavior, animal behavior and comparative psychology. Founded 1925.

Psychological Abstracts—Ohio State University, Columbus, Ohio: American Psychological Association. Subscription $7.00 a year ($7.25 foreign). 700 pages ann. Edited by W. S. Hunter. Monthly. Abstracts of psychological literature. Founded 1927.

Journal of Genetic Psychology—Provincetown, Mass.: The Journal Press. Subscription $14.00 per annum; $7.00 per volume. 1,000 pages annually (2 volumes). Edited by Carl Murchison and an international cooperating board. Quarterly. Child behavior, animal behavior, and comparative psychology. Founded 1891.

Journal of Social Psychology—Provincetown, Mass.: The Journal Press. Subscription $7.00. 500–600 pages annually. Edited by John Dewey, Carl Murchison, and an international cooperating board. Quarterly. Political, racial, and differential psychology. Founded 1929.

Personnel Journal—Baltimore: The Williams & Wilkins Co. Subscription $5.00. 500 pages, approximately. Edited by Walter V. Bingham. Bi-monthly.

Mental Measurement Monographs—Baltimore: The Williams & Wilkins Company. Subscription $5.00. 400 pages annually. Edited by Knight Dunlap and Buford Johnson. Published without fixed dates, each number a single research. Founded 1925.

Studies in Psychology and Psychiatry—Baltimore: The Williams & Wilkins Company. Subscription $5.00. 400 pages annually. Edited by Edward A. Pace. Published without fixed dates, each number a single research. Founded 1926.

Journal of Educational Psychology—Baltimore, Md.: Warwick & York. Subscription $6.00. Edited by Jack W. Dunlap, Harold E. Jones, and Percival M. Symonds. Published monthly except June to August. Founded 1910.

Child Development—Baltimore, Md.: Williams & Wilkins Company. Subscription $5.00. 400 pages annually. Edited by Buford Johnson. Quarterly. Founded 1930.

Character and Personality—Durham, N. C.: Duke University Press. Subscription $2.00 per annum. Quarterly. Published simultaneously in Great Britain and the United States. Edited by Charles E. Spearman (London) and Karl Zener (Duke University). Founded 1932.

Journal of Psychology—Provincetown, Mass.: The Journal Press. Subscription $14.00 per annum; $7.00 per volume. 1000–1200 pages annually. Immediate publication for any accepted manuscript. Edited by Carl Murchison and an editorial board. Quarterly. The general field of psychology. Founded 1935.

Directory of American Psychological Periodicals

American Journal of Psychology—Ithaca, N. Y.: Cornell University. Subscription $6.00. 600 pages annually. Edited by E. M. Dallenbach, M. F. Washburn, K. Dunlap, J. G. Beebe. Quarterly. General and experimental psychology. Founded 1887.

Journal of Abnormal Psychology—Provincetown, Mass.: The Journal Press. Subscription $4.00 per annum; $7.00 per volume. 500 pages annually (2 volumes). Edited by Carl Murchison and an international cooperating board. Quarterly. Experimental, theoretical, clinical, and historical psychology. Founded 1906.

Psychological Review—Ohio State University, Columbus, Ohio: Psychological Review Company. Subscription $5.00 (Foreign, $5.75). Single copies, $1.00. 540 pages annually. Bi-monthly. Original contributions only. Edited by Herbert S. Langfeld. Founded 1894.

Psychological Monographs—Ohio State University, Columbus, Ohio: Psychological Review Company. Subscription $6.00 per vol. 500 pages. Edited by John F. Dashiell. Published irregularly, each issue one or more researches. Founded 1896.

Psychological Index—Ohio State University, Columbus, Ohio: Psychological Review Company. Subscription $4.00. 300–400 pp. Edited by W. S. Hunter and R. R. Willoughby. An annual bibliography of psychological literature. Founded 1895.

Psychological Bulletin—Ohio State University, Columbus, Ohio: Psychological Review Company. Subscription $6.00. 700 pages annually. Psychological literature. Monthly (10 numbers). Edited by John A. McGeoch. Founded 1904.

Archives of Psychology—New York: Columbia University. Subscription $6.00. 500 pp. ann. Ed. by R. S. Woodworth. Published without fixed dates, each number a single experimental study. Founded 1906.

Journal of Abnormal and Social Psychology—Ohio State University, Columbus, Ohio: American Psychological Association. Subscription $5.00. Foreign, $5.25. Editor, Gordon W. Allport. Quarterly. 500 pages annually. Abnormal and Social. Founded 1906.

Psychological Clinic—Philadelphia: Psychological Clinic Press. Subscription $3.00. 288 pages. Ed. by Lightner Witmer. Without fixed dates (Quarterly). Orthogenics, psychology, hygiene. Founded 1907.

Psychoanalytic Review—Washington, D. C.: 3617 10th St., N.W. Subscription $5.50. 500 pages annually. Psychoanalysis. Quarterly. Ed. by W. A. White and S. E. Jelliffe. Founded 1913.

Journal of Experimental Psychology—Ohio State University, Columbus, Ohio. Monthly; two volumes per year, each volume of six numbers containing about 600 pages. Experimental. Subscription $14.00 ($7.00 per volume). Edited by H. W. Nemberger. Founded 1916.

Journal of Applied Psychology—Lancaster, Pa.: Science Press Printing Co. Subscription $6.00. 400 pages annually. Bi-monthly. Edited by James P. Porter, Ohio University, Athens, Ohio. Founded 1917.

(Continued on Third Cover)